(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
27 December 2001 (27.12.2001)

PCT

(10) International Publication Number
WO 01/98353 A2

(51) International Patent Classification[7]: C07K 14/705

(21) International Application Number: PCT/US01/19862

(22) International Filing Date: 20 June 2001 (20.06.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/212,890	20 June 2000 (20.06.2000)	US
60/213,466	23 June 2000 (23.06.2000)	US
60/214,601	27 June 2000 (27.06.2000)	US
60/222,372	31 July 2000 (31.07.2000)	US
60/231,435	8 September 2000 (08.09.2000)	US
60/232,889	15 September 2000 (15.09.2000)	US

(71) Applicant *(for all designated States except US)*: **INCYTE GENOMICS, INC.** [US/US]; 3160 Porter Drive, Palo Alto, CA 94304 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **HILLMAN, Jennifer, L.** [US/US]; 230 Monroe Drive #17, Mountain View, CA 94040 (US). **TANG, Y., Tom** [US/US]; 4230 Ranwick Court, San Jose, CA 95118 (US). **YUE, Henry** [US/US]; 826 Lois Avenue, Sunnyvale, CA 94087 (US). **ELLIOTT, Vicki, S.** [US/US]; 3770 Polton Place Way, San Jose, CA 95121 (US). **TRIBOULEY, Catherine, M.** [FR/US]; 1121 Tennessee Street #5, San Francisco, CA 94107 (US). **LEE, Ernestine, A.** [US/US]; 624 Kains Street, Albany, CA 94706 (US). **RAMKUMAR, Jayalaxmi** [IN/US]; 34359 Maybird Circle, Fremont, CA 94555 (US). **LAL, Preeti** [IN/US]; P.O. Box 5142, Santa Clara, CA 95056 (US). **XU, Yuming** [US/US]; 1739 Walnut Drive, Mountain View, CA 94040 (US). **WARREN, Bridget, A.** [US/US]; 10130 Parwood Drive #2, Cupertino, CA 95014 (US). **HAFALIA, April, J., A.** [US/US]; 2227 Calle de Primavera, Santa Clara, CA 95054 (US). **BAUGHN, Mariah, R.** [US/US]; 14244 Santiago Road, San Leandro, CA 94577 (US). **AZIMZAI, Yalda** [US/US]; 5518 Boulder Canyon Drive, Castro Valley, CA 94552 (US). **BATRA, Sajeev** [US/US]; 555 El Camino Real #709, Sunnyvale, CA 94087 (US). **BURFORD, Neil** [GB/US]; 105 Wildwood Circle, Durham, CT 06422 (US). **YAO, Monique, G.** [US/US]; 111 Frederick Court, Mountain View, CA 94043 (US). **NGUYEN, Danniel, B.** [US/US]; 1403 Ridgewood Drive, San Jose, CA 95118 (US). **LU, Dyung, Aina, M.** [US/US]; 233 Coy Drive, San Jose, CA 95123 (US). **WALIA, Narinder, K.** [US/US]; 890 Davis Street #205, San Leandro, CA 94577 (US). **GANDHI, Ameena, R.** [US/US]; 837 Roble Avenue #1, Menlo Park, CA 94025 (US). **AU-YOUNG, Janice** [US/US]; 233 Golden Eagle Lane, Brisbane, CA 94005 (US). **PATTERSON, Chandra** [US/US]; 490 Sherwood Way #1, Menlo Park, CA 94025 (US).

(74) Agents: **HAMLET-COX, Diana** et al.; Incyte Genomics, Inc., 3160 Porter Drive, Palo Alto, CA 94304 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

— *without international search report and to be republished upon receipt of that report*

— *with sequence listing part of description published separately in electronic form and available upon request from the International Bureau*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 01/98353 A2

(54) Title: SECRETED PROTEINS

(57) Abstract: The invention provides human secreted proteins (SECP) and polynucleotides which identify and encode SECP. The invention also provides expression vectors, host cells, antibodies, agonists, and antagonists. The invention also provides methods for diagnosing, treating, or preventing disorders associated with aberrant expression of SECP.

SECRETED PROTEINS

TECHNICAL FIELD

This invention relates to nucleic acid and amino acid sequences of secreted proteins and to the use of these sequences in the diagnosis, treatment, and prevention of cell proliferative, autoimmune/inflammatory, cardiovascular, neurological, and developmental disorders, and in the assessment of the effects of exogenous compounds on the expression of nucleic acid and amino acid sequences of secreted proteins.

BACKGROUND OF THE INVENTION

Protein transport and secretion are essential for cellular function. Protein transport is mediated by a signal peptide located at the amino terminus of the protein to be transported or secreted. The signal peptide is comprised of about ten to twenty hydrophobic amino acids which target the nascent protein from the ribosome to a particular membrane bound compartment such as the endoplasmic reticulum (ER). Proteins targeted to the ER may either proceed through the secretory pathway or remain in any of the secretory organelles such as the ER, Golgi apparatus, or lysosomes. Proteins that transit through the secretory pathway are either secreted into the extracellular space or retained in the plasma membrane. Proteins that are retained in the plasma membrane contain one or more transmembrane domains, each comprised of about 20 hydrophobic amino acid residues. Secreted proteins are generally synthesized as inactive precursors that are activated by post-translational processing events during transit through the secretory pathway. Such events include glycosylation, proteolysis, and removal of the signal peptide by a signal peptidase. Other events that may occur during protein transport include chaperone-dependent unfolding and folding of the nascent protein and interaction of the protein with a receptor or pore complex. Examples of secreted proteins with amino terminal signal peptides are discussed below and include proteins with important roles in cell-to-cell signaling. Such proteins include transmembrane receptors and cell surface markers, extracellular matrix molecules, cytokines, hormones, growth and differentiation factors, enzymes, neuropeptides, vasomediators, cell surface markers, and antigen recognition molecules. (Reviewed in Alberts, B. et al. (1994) Molecular Biology of The Cell, Garland Publishing, New York, NY, pp. 557-560, 582-592.)

Cell surface markers include cell surface antigens identified on leukocytic cells of the immune system. These antigens have been identified using systematic, monoclonal antibody (mAb)-based "shot gun" techniques. These techniques have resulted in the production of hundreds of mAbs directed against unknown cell surface leukocytic antigens. These antigens have been grouped into "clusters of differentiation" based on common immunocytochemical localization patterns in various

differentiated and undifferentiated leukocytic cell types. Antigens in a given cluster are presumed to identify a single cell surface protein and are assigned a "cluster of differentiation" or "CD" designation. Some of the genes encoding proteins identified by CD antigens have been cloned and verified by standard molecular biology techniques. CD antigens have been characterized as both transmembrane proteins and cell surface proteins anchored to the plasma membrane via covalent attachment to fatty acid-containing glycolipids such as glycosylphosphatidylinositol (GPI). (Reviewed in Barclay, A. N. et al. (1995) The Leucocyte Antigen Facts Book, Academic Press, San Diego, CA, pp. 17-20.)

Matrix proteins (MPs) are transmembrane and extracellular proteins which function in formation, growth, remodeling, and maintenance of tissues and as important mediators and regulators of the inflammatory response. The expression and balance of MPs may be perturbed by biochemical changes that result from congenital, epigenetic, or infectious diseases. In addition, MPs affect leukocyte migration, proliferation, differentiation, and activation in the immune response. MPs are frequently characterized by the presence of one or more domains which may include collagen-like domains, EGF-like domains, immunoglobulin-like domains, and fibronectin-like domains. In addition, MPs may be heavily glycosylated and may contain an Arginine-Glycine-Aspartate (RGD) tripeptide motif which may play a role in adhesive interactions. MPs include extracellular proteins such as fibronectin, collagen, galectin, vitronectin and its proteolytic derivative somatomedin B; and cell adhesion receptors such as cell adhesion molecules (CAMs), cadherins, and integrins. (Reviewed in Ayad, S. et al. (1994) The Extracellular Matrix Facts Book, Academic Press, San Diego, CA, pp. 2-16; Ruoslahti, E. (1997) Kidney Int. 51:1413-1417; Sjaastad, M.D. and Nelson, W.J. (1997) BioEssays 19:47-55.)

Mucins are highly glycosylated glycoproteins that are the major structural component of the mucus gel. The physiological functions of mucins are cytoprotection, mechanical protection, maintenance of viscosity in secretions, and cellular recognition. MUC6 is a human gastric mucin that is also found in gall bladder, pancreas, seminal vesicles, and female reproductive tract (Toribara, N.W. et al. (1997) J. Biol. Chem. 272:16398-16403). The MUC6 gene has been mapped to human chromosome 11 (Toribara, N.W. et al. (1993) J. Biol. Chem. 268:5879-5885). Hemomucin is a novel Drosophila surface mucin that may be involved in the induction of antibacterial effector molecules (Theopold, U. et al. (1996) J. Biol. Chem. 217:12708-12715).

Tuftelins are one of four different enamel matrix proteins that have been identified so far. The other three known enamel matrix proteins are the amelogenins, enamelin and ameloblastin. Assembly of the enamel extracellular matrix from these component proteins is believed to be critical in producing a matrix competent to undergo mineral replacement. (Paine C.T. et al. (1998) Connect Tissue Res.38:257-267). Tuftelin mRNA has been found to be expressed in human ameloblastoma

tumor, a non-mineralized odontogenic tumor (Deutsch D. et al. (1998) Connect Tissue Res. 39:177-184).

Olfactomedin-related proteins are extracellular matrix, secreted glycoproteins with conserved C-terminal motifs. They are expressed in a wide variety of tissues and in broad range of species, from *Caenorhabditis elegans* to *Homo sapiens*. Olfactomedin-related proteins comprise a gene family with at least 5 family members in humans. One of the five, TIGR/myocilin protein, is expressed in the eye and is associated with the pathogenesis of glaucoma (Kulkarni, N.H. et al., (2000) Genet. Res. 76:41-50). Research by Yokoyama et al. (1996) found a 135-amino acid protein, termed AMY, having 96% sequence identity with rat neuronal olfactomedin-releated ER localized protein in a neuroblastoma cell line cDNA library, suggesting an essential role for AMY in nerve tissue (Yokoyama, M. et al., (1996) DNA Res. 3:311-320). Neuron-specific olfactomedin-related glycoproteins isolated from rat brain cDNA libraries show strong sequence similarity with olfactomedin. This similarity is suggestive of a matrix-related function of these glycoproteins in neurons and neurosecretory cells (Danielson, P.E. et al., (1994) J. Neurosci. Res. 38:468-478).

Mac-2 binding protein is a 90-kD serum protein (90K) and another secreted glycoprotein, isolated from both the human breast carcinoma cell line SK-BR-3, and human breast milk. It specifically binds to a human macrophage-associated lectin, Mac-2. Structurally, the mature protein is 567 amino acids in length and is proceeded by an 18-amino acid leader. There are 16 cysteines and seven potential N-linked glycosylation sites. The first 106 amino acids represent a domain very similar to an ancient protein superfamily defined by a macrophage scavenger receptor cysteine-rich domain (Koths,K. et al., (1993) J. Biol. Chem. 268:14245-14249). 90K is elevated in the serum of subpopulations of AIDS patients and is expressed at varying levels in primary tumor samples and tumor cell lines. Ullrich et al. (1994) have demonstrated that 90K stimulates host defense systems and can induce interleukin-2 secretion. This immune stimulation is proposed to be a result of oncogenic transformation, viral infection or pathogenic invasion (Ullrich,A., et al. (1994) J. Biol. Chem. 269:18401-18407).

Semaphorins are a large group of axonal guidance molecules consisting of at least 30 different members and are found in vertebrates, invertebrates, and even certain viruses. All semaphorins contain the sema domain which is approximately 500 amino acids in length. Neuropilin, a semaphorin receptor has been shown to promote neurite outgrowth in vitro. The extracellular region of neuropilins consists of three different domains: CUB, discoidin, and MAM domains. The CUB and the MAM motifs of neuropilin have been suggested as having roles in protein-protein interactions and are suggested to be involved in the binding of semaphorins through the sema and the C-terminal domains (reviewed in Raper, J.A. (2000) Curr. Opin. Neurobiol. 10:88-94). Plexins are neuronal cell surface molecules that mediate cell adhesion via a homophilic binding mechanism in the

presence of calcium ions. Plexins have been shown to be expressed in the receptors and neurons of particular sensory systems (Ohta, K. et al. (1995) Cell 14:1189-1199). There is evidence that suggests that some plexins function to control motor and CNS axon guidance in the developing nervous system. Plexins, which themselves contain complete semaphorin domains, may be both the ancestors of classical semaphorins and binding partners for semaphorins (Winberg, M.L. et al (1998) Cell 95:903-916).

Human pregnancy-specific beta 1-glycoprotein (PSG) is a family of closely related glycoproteins of molecular weights of 72 KDa, 64KDa, 62KDa, and 54KDa. Together with the carcinoembryonic antigen, they comprise a subfamily within the immunoglobulin superfamily (Plouzek C.A. and Chou J.Y., Endocrinology 129:950-958) Different subpopulations of PSG have been found to be produced by the trophoblasts of the human placenta, and the amnionic, and chorionic membranes (Plouzek C.A. et al. (1993) Placenta 14:277-285).

Autocrine motility factor (AMF) is one of the motility cytokines regulating tumor cell migration, therefore identification of the signaling pathway coupled with it has critical importance. Autocrine motility factor receptor (AMFR) expression has been found to be associated with tumor progression in thymoma (Ohta Y. et al. (2000) Int. J. Oncol. 17:259-264). AMFR is a cell surface glycoprotein of molecular weight 78KDa.

Hormones are secreted molecules that travel through the circulation and bind to specific receptors on the surface of, or within, target cells. Although they have diverse biochemical compositions and mechanisms of action, hormones can be grouped into two categories. One category includes small lipophilic hormones that diffuse through the plasma membrane of target cells, bind to cytosolic or nuclear receptors, and form a complex that alters gene expression. Examples of these molecules include retinoic acid, thyroxine, and the cholesterol-derived steroid hormones such as progesterone, estrogen, testosterone, cortisol, and aldosterone. The second category includes hydrophilic hormones that function by binding to cell surface receptors that transduce signals across the plasma membrane. Examples of such hormones include amino acid derivatives such as catecholamines (epinephrine, norepinephrine) and histamine, and peptide hormones such as glucagon, insulin, gastrin, secretin, cholecystokinin, adrenocorticotropic hormone, follicle stimulating hormone, luteinizing hormone, thyroid stimulating hormone, and vasopressin. (See, for example, Lodish et al. (1995) Molecular Cell Biology, Scientific American Books Inc., New York, NY, pp. 856-864.)

Pro-opiomelanocortin (POMC) is the precursor polypeptide of corticotropin (ACTH) a hormone synthesized by the anterior pituitary gland, which functions in the stimulation of the adrenal cortex. POMC is also the precursor polypeptide of the hormone, beta-lipotropin (beta-LPH),. Each hormone includes smaller peptides with distinct biological activities: alpha-melanotropin (alpha-MSH) and corticotropin-like intermediate lobe peptide (CLIP) are formed from ACTH; gamma-

lipotropin (gamma-LPH) and beta-endorphin are peptide components of beta-LPH, while beta-MSH is contained within gamma-LPH. Adrenal insufficiency due to ACTH deficiency, resulting from a genetic mutation in exons 2 and 3 of POMC results in an endocrine disorder characterized by early-onset obesity, adrenal insufficiency, and red hair pigmentation (Chretien, M. et al., (1979) Canad. J. Biochem. 57:1111-1121, Krude, H. et al., (1998) Nature Genet. 19:155-157, Online Mendelian Inheritance in Man, OMIM. Johns Hopkins University, Baltimore, MD. OMIM Number: 176830: August 1, 2000. World Wide Web URL: www.ncbi.nlm.nih.gov/omim/).

Growth and differentiation factors are secreted proteins which function in intercellular communication. Some factors require oligomerization or association with membrane proteins for activity. Complex interactions among these factors and their receptors trigger intracellular signal transduction pathways that stimulate or inhibit cell division, cell differentiation, cell signaling, and cell motility. Most growth and differentiation factors act on cells in their local environment (paracrine signaling). There are three broad classes of growth and differentiation factors. The first class includes the large polypeptide growth factors such as epidermal growth factor, fibroblast growth factor, transforming growth factor, insulin-like growth factor, and platelet-derived growth factor. The second class includes the hematopoietic growth factors such as the colony stimulating factors (CSFs). Hematopoietic growth factors stimulate the proliferation and differentiation of blood cells such as B-lymphocytes, T-lymphocytes, erythrocytes, platelets, eosinophils, basophils, neutrophils, macrophages, and their stem cell precursors. The third class includes small peptide factors such as bombesin, vasopressin, oxytocin, endothelin, transferrin, angiotensin II, vasoactive intestinal peptide, and bradykinin which function as hormones to regulate cellular functions other than proliferation.

Growth and differentiation factors play critical roles in neoplastic transformation of cells in vitro and in tumor progression in vivo. Inappropriate expression of growth factors by tumor cells may contribute to vascularization and metastasis of tumors. During hematopoiesis, growth factor misregulation can result in anemias, leukemias, and lymphomas. Certain growth factors such as interferon are cytotoxic to tumor cells both in vivo and in vitro. Moreover, some growth factors and growth factor receptors are related both structurally and functionally to oncoproteins. In addition, growth factors affect transcriptional regulation of both proto-oncogenes and oncosuppressor genes. (Reviewed in Pimentel, E. (1994) Handbook of Growth Factors, CRC Press, Ann Arbor, MI, pp. 1-9.)

The Slit protein, first identified in Drosophila, is critical in central nervous system midline formation and potentially in nervous tissue histogenesis and axonal pathfinding. Itoh et al. have identified mammalian homologues of the slit gene (human Slit-1, Slit-2, Slit-3 and rat Slit-1). The encoded proteins are putative secreted proteins containing EFG-like motifs and leucine-rich repeats, both are conserved protein-protein interaction domains. Slit-1, -2, and -3 mRNAs are expressed in the brain, spinal cord, and thyroid, respectively (Itoh, A. et al., (1998) Brain Res. Mol. Brain Res.

62:175-186). The Slit family of proteins are indicated to be functional ligands of glypican-1 in nervous tissue and suggests that their interactions may be critical in certain stages during central nervous system histogenesis (Liang, Y. et al., (1999) J. Biol. Chem. 274:17885-17892).

Neuropeptides and vasomediators (NP/VM) comprise a large family of endogenous signaling molecules. Included in this family are neuropeptides and neuropeptide hormones such as bombesin, neuropeptide Y, neurotensin, neuromedin N, melanocortins, opioids, galanin, somatostatin, tachykinins, urotensin II and related peptides involved in smooth muscle stimulation, vasopressin, vasoactive intestinal peptide, and circulatory system-borne signaling molecules such as angiotensin, complement, calcitonin, endothelins, formyl-methionyl peptides, glucagon, cholecystokinin and gastrin. NP/VMs can transduce signals directly, modulate the activity or release of other neurotransmitters and hormones, and act as catalytic enzymes in cascades. The effects of NP/VMs range from extremely brief to long-lasting. (Reviewed in Martin, C.R. et al. (1985) Endocrine Physiology, Oxford University Press, New York, NY, pp. 57-62.)

NP/VMs are involved in numerous neurological and cardiovascular disorders. For example, neuropeptide Y is involved in hypertension, congestive heart failure, affective disorders, and appetite regulation. Somatostatin inhibits secretion of growth hormone and prolactin in the anterior pituitary, as well as inhibiting secretion in intestine, pancreatic acinar cells, and pancreatic beta-cells. A reduction in somatostatin levels has been reported in Alzheimer's disease and Parkinson's disease. Vasopressin acts in the kidney to increase water and sodium absorption, and in higher concentrations stimulates contraction of vascular smooth muscle, platelet activation, and glycogen breakdown in the liver. Vasopressin and its analogues are used clinically to treat diabetes insipidus. Endothelin and angiotensin are involved in hypertension, and drugs, such as captopril, which reduce plasma levels of angiotensin, are used to reduce blood pressure (Watson, S. and S. Arkinstall (1994) The G-protein Linked Receptor Facts Book, Academic Press, San Diego CA, pp. 194; 252; 284; 55; 111).

Neuropeptides have also been shown to have roles in nociception (pain). Vasoactive intestinal peptide appears to play an important role in chronic neuropathic pain. Nociceptin, an endogenous ligand for for the opioid receptor-like 1 receptor, is thought to have a predominantly anti-nociceptive effect, and has been shown to have analgesic properties in different animal models of tonic or chronic pain (Dickinson, T. and Fleetwood-Walker, S.M. (1998) Trends Pharmacol. Sci. 19:346-348).

Other proteins that contain signal peptides include secreted proteins with enzymatic activity. Such activity includes, for example, oxidoreductase/dehydrogenase activity, transferase activity, hydrolase activity, lyase activity, isomerase activity, or ligase activity. For example, matrix metalloproteinases are secreted hydrolytic enzymes that degrade the extracellular matrix and thus play an important role in tumor metastasis, tissue morphogenesis, and arthritis (Reponen, P. et al.

(1995) Dev. Dyn. 202:388-396; Firestein, G.S. (1992) Curr. Opin. Rheumatol. 4:348-354; Ray, J.M. and Stetler-Stevenson, W.G. (1994) Eur. Respir. J. 7:2062-2072; and Mignatti, P. and Rifkin, D.B. (1993) Physiol. Rev. 73:161-195). Additional examples are the acetyl-CoA synthetases which activate acetate for use in lipid synthesis or energy generation (Luong, A. et al. (2000) J. Biol. Chem. 275:26458-26466). The result of acetyl-CoA synthetase activity is the formation of acetyl-CoA from acetate and CoA. Acetyl-CoA sythetases share a region of sequence similarity identified as the AMP-binding domain signature. Acetyl-CoA synthetase has been shown to be associated with hypertension (H. Toh (1991) Protein Seq. Data Anal. 4:111-117 and Iwai, N. et al., (1994) Hypertension 23:375-380).

Other proteins that contain signal peptides include enzymes involved in the glycosylation of proteins in transit through the secretory pathway. Mucin-type O-linked glycosylation is a dominant form of protein glycosylation. Initiation of mucin-type glycosylation occurs by the addition of the monosaccharide N-acetylgalactosamine to the hydroxyl group of serine and threonine amino acids (GalNAcα1-O-Ser/Thr). GalNAc O-glycosylation is more prominent on high molecular weight secretory glycoproteins such as mucins, but is also found on a variety of glycoproteins (White, T. et. al., J. Biol. Chem. (1995) 270:24156-24165). Additionally, serine/threonine-rich tandem repeats are a characteristic of human mucin core proteins. The tandem repeat region also contains numerous antigenic determinants as recognized by the monoclonal antibodies HMFG-1, HMFG-1, and SM-3. Glycosylation sites within the tandem repeat region were found to be differentially glycosylated depending on the organ from which Muc1 was isolated. The finding of variable glycosylation activity may be critical to further understanding of the molecular basis of cancer-associated epitopes which map to the Muc1 tandem repeat (Gendler, S.J. et al. (1990) J. Biol. Chem. 265:15286-15293).

Antigen recognition molecules are key players in the sophisticated and complex immune systems which all vertebrates have developed to provide protection from viral, bacterial, fungal, and parasitic infections. A key feature of the immune system is its ability to distinguish foreign molecules, or antigens, from "self" molecules. This ability is mediated primarily by secreted and transmembrane proteins expressed by leukocytes (white blood cells) such as lymphocytes, granulocytes, and monocytes. Most of these proteins belong to the immunoglobulin (Ig) superfamily, members of which contain one or more repeats of a conserved structural domain. This Ig domain is comprised of antiparallel β sheets joined by a disulfide bond in an arrangement called the Ig fold. Members of the Ig superfamily include T-cell receptors, major histocompatibility (MHC) proteins, antibodies, and immune cell-specific surface markers such as the "cluster of differentiation" or CD antigens. These antigens have been identified using systematic, monoclonal antibody (mAb)-based "shot gun" techniques. These techniques have resulted in the production of hundreds of mAbs directed against unknown cell surface leukocytic antigens. These antigens have been grouped into

"clusters of differentiation" based on common immunocytochemical localization patterns in various differentiated and undifferentiated leukocytic cell types. Antigens in a given cluster are presumed to identify a single cell surface protein and are assigned a "cluster of differentiation" or "CD" designation. Some of the genes encoding proteins identified by CD antigens have been cloned and verified by standard molecular biology techniques. CD antigens have been characterized as both transmembrane proteins and cell surface proteins anchored to the plasma membrane via covalent attachment to fatty acid-containing glycolipids such as glycosylphosphatidylinositol (GPI). (Reviewed in Barclay, A. N. et al. (1995) The Leucocyte Antigen Facts Book, Academic Press, San Diego, CA, pp. 17-20.)

MHC proteins are cell surface markers that bind to and present foreign antigens to T cells. MHC molecules are classified as either class I or class II. Class I MHC molecules (MHC I) are expressed on the surface of almost all cells and are involved in the presentation of antigen to cytotoxic T cells. For example, a cell infected with virus will degrade intracellular viral proteins and express the protein fragments bound to MHC I molecules on the cell surface. The MHC I/antigen complex is recognized by cytotoxic T-cells which destroy the infected cell and the virus within. Class II MHC molecules are expressed primarily on specialized antigen-presenting cells of the immune system, such as B-cells and macrophages. These cells ingest foreign proteins from the extracellular fluid and express MHC II/antigen complex on the cell surface. This complex activates helper T-cells, which then secrete cytokines and other factors that stimulate the immune response. MHC molecules also play an important role in organ rejection following transplantation. Rejection occurs when the recipient's T-cells respond to foreign MHC molecules on the transplanted organ in the same way as to self MHC molecules bound to foreign antigen. (Reviewed in Alberts, B. et al. (1994) Molecular Biology of the Cell, Garland Publishing, New York, NY, pp. 1229-1246.)

Antibodies, or immunoglobulins, are either expressed on the surface of B-cells or secreted by B-cells into the circulation. Antibodies bind and neutralize foreign antigens in the blood and other extracellular fluids. The prototypical antibody is a tetramer consisting of two identical heavy polypeptide chains (H-chains) and two identical light polypeptide chains (L-chains) interlinked by disulfide bonds. This arrangement confers the characteristic Y-shape to antibody molecules. Antibodies are classified based on their H-chain composition. The five antibody classes, IgA, IgD, IgE, IgG and IgM, are defined by the α, δ, ϵ, γ, and μ H-chain types. There are two types of L-chains, κ and λ, either of which may associate as a pair with any H-chain pair. IgG, the most common class of antibody found in the circulation, is tetrameric, while the other classes of antibodies are generally variants or multimers of this basic structure.

H-chains and L-chains each contain an N-terminal variable region and a C-terminal constant region. The constant region consists of about 110 amino acids in L-chains and about 330 or 440

amino acids in H-chains. The amino acid sequence of the constant region is nearly identical among H- or L-chains of a particular class. The variable region consists of about 110 amino acids in both H- and L-chains. However, the amino acid sequence of the variable region differs among H- or L-chains of a particular class. Within each H- or L-chain variable region are three hypervariable regions of extensive sequence diversity, each consisting of about 5 to 10 amino acids. In the antibody molecule, the H- and L-chain hypervariable regions come together to form the antigen recognition site. (Reviewed in Alberts, supra, pp. 1206-1213 and 1216-1217.)

Both H-chains and L-chains contain repeated Ig domains. For example, a typical H-chain contains four Ig domains, three of which occur within the constant region and one of which occurs within the variable region and contributes to the formation of the antigen recognition site. Likewise, a typical L-chain contains two Ig domains, one of which occurs within the constant region and one of which occurs within the variable region.

The immune system is capable of recognizing and responding to any foreign molecule that enters the body. Therefore, the immune system must be armed with a full repertoire of antibodies against all potential antigens. Such antibody diversity is generated by somatic rearrangement of gene segments encoding variable and constant regions. These gene segments are joined together by site-specific recombination which occurs between highly conserved DNA sequences that flank each gene segment. Because there are hundreds of different gene segments, millions of unique genes can be generated combinatorially. In addition, imprecise joining of these segments and an unusually high rate of somatic mutation within these segments further contribute to the generation of a diverse antibody population.

A number of isomerases catalyze steps in protein folding, phototransduction, and various anabolic and catabolic pathways. One class of isomerases is known as peptidyl-prolyl *cis-trans* isomerases (PPIases). PPIases catalyze the *cis* to *trans* isomerization of certain proline imidic bonds in proteins. Two families of PPIases are the FK506 binding proteins (FKBPs), and cyclophilins (CyPs). FKBPs bind the potent immunosuppressants FK506 and rapamycin, thereby inhibiting signaling pathways in T-cells. Specifically, the PPIase activity of FKBPs is inhibited by binding of FK506 or rapamycin. There are five members of the FKBP family which are named according to their calculated molecular masses (FKBP12, FKBP13, FKBP25, FKBP52, and FKBP65), and localized to different regions of the cell where they associate with different protein complexes (Coss, M. et al. (1995) J. Biol. Chem. 270:29336 - 29341; Schreiber, S.L. (1991) Science 251:283 - 287).

The peptidyl-prolyl isomerase activity of CyP may be part of the signaling pathway that leads to T-cell activation. CyP isomerase activity is associated with protein folding and protein trafficking, and may also be involved in assembly/disassembly of protein complexes and regulation of protein activity. For example, in *Drosophila*, the CyP NinaA is required for correct localization of

rhodopsins, while a mammalian CyP (Cyp40) is part of the Hsp90/Hsc70 complex that binds steroid receptors. The mammalian CypA has been shown to bind the *gag* protein from human immunodeficiency virus 1 (HIV-1), an interaction that can be inhibited by cyclosporin. Since cyclosporin has potent anti-HIV-1 activity, CypA may play an essential function in HIV-1 replication. Finally, Cyp40 has been shown to bind and inactivate the transcription factor c-Myb, an effect that is reversed by cyclosporin. This effect implicates CyPs in the regulation of transcription, transformation, and differentiation (Bergsma, D.J. et al (1991) J. Biol. Chem. 266:23204 - 23214; Hunter, T. (1998) Cell 92: 141-143; and Leverson, J.D. and Ness, S.A. (1998) Mol. Cell. 1:203-211).

Gamma-carboxyglutamic acid (Gla) proteins rich in proline (PRGPs) are members of a family of vitamin K-dependent single-pass integral membrane proteins. These proteins are characterized by an extracellular amino terminal domain of approximately 45 amino acids rich in Gla. The intracellular carboxyl terminal region contains one or two copies of the sequence PPXY, a motif present in a variety of proteins involved in such diverse cellular functions as signal transduction, cell cycle progression, and protein turnover (Kulman, J.D. et al., (2001) Proc. Natl. Acad. Sci. U.S.A. 98:1370-1375). The process of post-translational modification of glutamic residues to form Gla is Vitamin K-dependent carboxylation. Proteins which contain Gla include plasma proteins involved in blood coagulation. These proteins are prothrombin, proteins C, S, and Z, and coagulation factors VII, IX, and X. Osteocalcin (bone-Gla protein, BGP) and matrix Gla-protein (MGP) also contain Gla (Friedman, P.A., and C.T. Przysiecki (1987) Int. J. Biochem. 19:1-7; C. Vermeer (1990) Biochem. J. 266:625-636).

The Drosophila sp. gene *crossveinless 2* is characterized as having a putative signal or transmembrane sequence, and a partial Von Willebrand Factor D domain similar to those domains known to regulate the formation of intramolecular and intermolecular bonds and five cysteine-rich domains, known to bind BMP-like (bone morphogenetic proteins) ligands. These features suggest that *crossveinless 2* may act extracelluarly or in the secretory pathway to directly potentiate ligand signaling and hence, involvement in the BMP-like signaling pathway known to play a role in vein specification (Conley, C.A. et al., (2000) Development 127:3947-3959). The dorsal-ventral patterning in both vertebrate and Drosophila embryos requires a conserved system of extracellular proteins to generate a positional informational gradient.

Another protein that contains a signal peptide is encoded by the seizure-related gene, SEZ-6, a brain specific cDNA whose expression is increased by the convulsant drug pentylentetrazole. The SEZ-6 protein is expressed in the cerebrum and cerebellum. SEZ-6 contains five short consensus repeats (SCR, or sushi domains) and two CUB (complement C1r/s-like repeat) domains in addition to a signal peptide and a single transmembrane domain (Shimizu-Nishikawa, K. et al. (1995) Biochem. Biophys. Res. Commun. 216:382-389).

The discovery of new secreted proteins and the polynucleotides encoding them satisfies a need in the art by providing new compositions which are useful in the diagnosis, prevention, and treatment of cell proliferative, autoimmune/inflammatory, cardiovascular, neurological, and developmental disorders, and in the assessment of the effects of exogenous compounds on the expression of nucleic acid and amino acid sequences of secreted proteins.

SUMMARY OF THE INVENTION

The invention features purified polypeptides, secreted proteins, referred to collectively as "SECP" and individually as "SECP-1," "SECP-2," "SECP-3," "SECP-4," "SECP-5," "SECP-6," "SECP-7," "SECP-8," "SECP-9," "SECP-10," "SECP-11," "SECP-12," "SECP-13," "SECP-14," "SECP-15," "SECP-16," "SECP-17," "SECP-18," "SECP-19," "SECP-20," "SECP-21," "SECP-22," "SECP-23," "SECP-24," "SECP-25," "SECP-26," "SECP-27," "SECP-28," "SECP-29," "SECP-30," "SECP-31," "SECP-32," "SECP-33," "SECP-34," "SECP-35," "SECP-36," "SECP-37," "SECP-38," "SECP-39," "SECP-40," "SECP-41," "SECP-42," "SECP-43," and "SECP-44." In one aspect, the invention provides an isolated polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-44, b) a naturally occurring polypeptide comprising an amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-44, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-44, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-44. In one alternative, the invention provides an isolated polypeptide comprising the amino acid sequence of SEQ ID NO:1-44.

The invention further provides an isolated polynucleotide encoding a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-44, b) a naturally occurring polypeptide comprising an amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-44, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-44, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-44. In one alternative, the polynucleotide encodes a polypeptide selected from the group consisting of SEQ ID NO:1-44. In another alternative, the polynucleotide is selected from the group consisting of SEQ ID NO:45-88.

Additionally, the invention provides a recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide encoding a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting

of SEQ ID NO:1-44, b) a naturally occurring polypeptide comprising an amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-44, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-44, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-44. In one alternative, the invention provides a cell transformed with the recombinant polynucleotide. In another alternative, the invention provides a transgenic organism comprising the recombinant polynucleotide.

The invention also provides a method for producing a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-44, b) a naturally occurring polypeptide comprising an amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-44, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-44, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-44. The method comprises a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell is transformed with a recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide encoding the polypeptide, and b) recovering the polypeptide so expressed.

Additionally, the invention provides an isolated antibody which specifically binds to a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-44, b) a naturally occurring polypeptide comprising an amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-44, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-44, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-44.

The invention further provides an isolated polynucleotide selected from the group consisting of a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:45-88, b) a naturally occurring polynucleotide comprising a polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:45-88, c) a polynucleotide complementary to the polynucleotide of a), d) a polynucleotide complementary to the polynucleotide of b), and e) an RNA equivalent of a)-d). In one alternative, the polynucleotide comprises at least 60 contiguous nucleotides.

Additionally, the invention provides a method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide selected from the group consisting of a) a polynucleotide comprising a polynucleotide sequence selected from the group

consisting of SEQ ID NO:45-88, b) a naturally occurring polynucleotide comprising a polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:45-88, c) a polynucleotide complementary to the polynucleotide of a), d) a polynucleotide complementary to the polynucleotide of b), and e) an RNA equivalent of a)-d). The method comprises a) hybridizing the sample with a probe comprising at least 20 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization complex is formed between said probe and said target polynucleotide or fragments thereof, and b) detecting the presence or absence of said hybridization complex, and optionally, if present, the amount thereof. In one alternative, the probe comprises at least 60 contiguous nucleotides.

The invention further provides a method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide selected from the group consisting of a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:45-88, b) a naturally occurring polynucleotide comprising a polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:45-88, c) a polynucleotide complementary to the polynucleotide of a), d) a polynucleotide complementary to the polynucleotide of b), and e) an RNA equivalent of a)-d). The method comprises a) amplifying said target polynucleotide or fragment thereof using polymerase chain reaction amplification, and b) detecting the presence or absence of said amplified target polynucleotide or fragment thereof, and, optionally, if present, the amount thereof.

The invention further provides a composition comprising an effective amount of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-44, b) a naturally occurring polypeptide comprising an amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-44, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-44, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-44, and a pharmaceutically acceptable excipient. In one embodiment, the composition comprises an amino acid sequence selected from the group consisting of SEQ ID NO:1-44. The invention additionally provides a method of treating a disease or condition associated with decreased expression of functional SECP, comprising administering to a patient in need of such treatment the composition.

The invention also provides a method for screening a compound for effectiveness as an agonist of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino

acid sequence selected from the group consisting of SEQ ID NO:1-44, b) a naturally occurring polypeptide comprising an amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-44, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-44, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-44. The method comprises a) exposing a sample comprising the polypeptide to a compound, and b) detecting agonist activity in the sample. In one alternative, the invention provides a composition comprising an agonist compound identified by the method and a pharmaceutically acceptable excipient. In another alternative, the invention provides a method of treating a disease or condition associated with decreased expression of functional SECP, comprising administering to a patient in need of such treatment the composition.

Additionally, the invention provides a method for screening a compound for effectiveness as an antagonist of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-44, b) a naturally occurring polypeptide comprising an amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-44, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-44, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-44. The method comprises a) exposing a sample comprising the polypeptide to a compound, and b) detecting antagonist activity in the sample. In one alternative, the invention provides a composition comprising an antagonist compound identified by the method and a pharmaceutically acceptable excipient. In another alternative, the invention provides a method of treating a disease or condition associated with overexpression of functional SECP, comprising administering to a patient in need of such treatment the composition.

The invention further provides a method of screening for a compound that specifically binds to a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-44, b) a naturally occurring polypeptide comprising an amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-44, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-44, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-44. The method comprises a) combining the polypeptide with at least one test compound under suitable conditions, and b) detecting binding of the polypeptide to the test compound, thereby identifying a compound that specifically binds to the polypeptide.

The invention further provides a method of screening for a compound that modulates the

activity of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-44, b) a naturally occurring polypeptide comprising an amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-44, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-44, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-44. The method comprises a) combining the polypeptide with at least one test compound under conditions permissive for the activity of the polypeptide, b) assessing the activity of the polypeptide in the presence of the test compound, and c) comparing the activity of the polypeptide in the presence of the test compound with the activity of the polypeptide in the absence of the test compound, wherein a change in the activity of the polypeptide in the presence of the test compound is indicative of a compound that modulates the activity of the polypeptide.

The invention further provides a method for screening a compound for effectiveness in altering expression of a target polynucleotide, wherein said target polynucleotide comprises a sequence selected from the group consisting of SEQ ID NO:45-88, the method comprising a) exposing a sample comprising the target polynucleotide to a compound, and b) detecting altered expression of the target polynucleotide.

The invention further provides a method for assessing toxicity of a test compound, said method comprising a) treating a biological sample containing nucleic acids with the test compound; b) hybridizing the nucleic acids of the treated biological sample with a probe comprising at least 20 contiguous nucleotides of a polynucleotide selected from the group consisting of i) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:45-88, ii) a naturally occurring polynucleotide comprising a polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:45-88, iii) a polynucleotide having a sequence complementary to i), iv) a polynucleotide complementary to the polynucleotide of ii), and v) an RNA equivalent of i)-iv). Hybridization occurs under conditions whereby a specific hybridization complex is formed between said probe and a target polynucleotide in the biological sample, said target polynucleotide selected from the group consisting of i) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:45-88, ii) a naturally occurring polynucleotide comprising a polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:45-88, iii) a polynucleotide complementary to the polynucleotide of i), iv) a polynucleotide complementary to the polynucleotide of ii), and v) an RNA equivalent of i)-iv). Alternatively, the target polynucleotide comprises a fragment of a polynucleotide sequence selected from the group consisting

of i)-v) above; c) quantifying the amount of hybridization complex; and d) comparing the amount of hybridization complex in the treated biological sample with the amount of hybridization complex in an untreated biological sample, wherein a difference in the amount of hybridization complex in the treated biological sample is indicative of toxicity of the test compound.

BRIEF DESCRIPTION OF THE TABLES

Table 1 summarizes the nomenclature for the full length polynucleotide and polypeptide sequences of the present invention.

Table 2 shows the GenBank identification number and annotation of the nearest GenBank homolog for polypeptides of the invention. The probability score for the match between each polypeptide and its GenBank homolog is also shown.

Table 3 shows structural features of polypeptide sequences of the invention, including predicted motifs and domains, along with the methods, algorithms, and searchable databases used for analysis of the polypeptides.

Table 4 lists the cDNA and/or genomic DNA fragments which were used to assemble polynucleotide sequences of the invention, along with selected fragments of the polynucleotide sequences.

Table 5 shows the representative cDNA library for polynucleotides of the invention.

Table 6 provides an appendix which describes the tissues and vectors used for construction of the cDNA libraries shown in Table 5.

Table 7 shows the tools, programs, and algorithms used to analyze the polynucleotides and polypeptides of the invention, along with applicable descriptions, references, and threshold parameters.

DESCRIPTION OF THE INVENTION

Before the present proteins, nucleotide sequences, and methods are described, it is understood that this invention is not limited to the particular machines, materials and methods described, as these may vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to limit the scope of the present invention which will be limited only by the appended claims.

It must be noted that as used herein and in the appended claims, the singular forms "a," "an," and "the" include plural reference unless the context clearly dictates otherwise. Thus, for example, a reference to "a host cell" includes a plurality of such host cells, and a reference to "an antibody" is a reference to one or more antibodies and equivalents thereof known to those skilled in the art, and so forth.

Unless defined otherwise, all technical and scientific terms used herein have the same meanings as commonly understood by one of ordinary skill in the art to which this invention belongs. Although any machines, materials, and methods similar or equivalent to those described herein can be used to practice or test the present invention, the preferred machines, materials and methods are now described. All publications mentioned herein are cited for the purpose of describing and disclosing the cell lines, protocols, reagents and vectors which are reported in the publications and which might be used in connection with the invention. Nothing herein is to be construed as an admission that the invention is not entitled to antedate such disclosure by virtue of prior invention.

DEFINITIONS

"SECP" refers to the amino acid sequences of substantially purified SECP obtained from any species, particularly a mammalian species, including bovine, ovine, porcine, murine, equine, and human, and from any source, whether natural, synthetic, semi-synthetic, or recombinant.

The term "agonist" refers to a molecule which intensifies or mimics the biological activity of SECP. Agonists may include proteins, nucleic acids, carbohydrates, small molecules, or any other compound or composition which modulates the activity of SECP either by directly interacting with SECP or by acting on components of the biological pathway in which SECP participates.

An "allelic variant" is an alternative form of the gene encoding SECP. Allelic variants may result from at least one mutation in the nucleic acid sequence and may result in altered mRNAs or in polypeptides whose structure or function may or may not be altered. A gene may have none, one, or many allelic variants of its naturally occurring form. Common mutational changes which give rise to allelic variants are generally ascribed to natural deletions, additions, or substitutions of nucleotides. Each of these types of changes may occur alone, or in combination with the others, one or more times in a given sequence.

"Altered" nucleic acid sequences encoding SECP include those sequences with deletions, insertions, or substitutions of different nucleotides, resulting in a polypeptide the same as SECP or a polypeptide with at least one functional characteristic of SECP. Included within this definition are polymorphisms which may or may not be readily detectable using a particular oligonucleotide probe of the polynucleotide encoding SECP, and improper or unexpected hybridization to allelic variants, with a locus other than the normal chromosomal locus for the polynucleotide sequence encoding SECP. The encoded protein may also be "altered," and may contain deletions, insertions, or substitutions of amino acid residues which produce a silent change and result in a functionally equivalent SECP. Deliberate amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues, as long as the biological or immunological activity of SECP is retained. For example, negatively charged amino acids may include aspartic acid and glutamic acid, and positively charged

amino acids may include lysine and arginine. Amino acids with uncharged polar side chains having similar hydrophilicity values may include: asparagine and glutamine; and serine and threonine. Amino acids with uncharged side chains having similar hydrophilicity values may include: leucine, isoleucine, and valine; glycine and alanine; and phenylalanine and tyrosine.

The terms "amino acid" and "amino acid sequence" refer to an oligopeptide, peptide, polypeptide, or protein sequence, or a fragment of any of these, and to naturally occurring or synthetic molecules. Where "amino acid sequence" is recited to refer to a sequence of a naturally occurring protein molecule, "amino acid sequence" and like terms are not meant to limit the amino acid sequence to the complete native amino acid sequence associated with the recited protein molecule.

"Amplification" relates to the production of additional copies of a nucleic acid sequence. Amplification is generally carried out using polymerase chain reaction (PCR) technologies well known in the art.

The term "antagonist" refers to a molecule which inhibits or attenuates the biological activity of SECP. Antagonists may include proteins such as antibodies, nucleic acids, carbohydrates, small molecules, or any other compound or composition which modulates the activity of SECP either by directly interacting with SECP or by acting on components of the biological pathway in which SECP participates.

The term "antibody" refers to intact immunoglobulin molecules as well as to fragments thereof, such as Fab, $F(ab')_2$, and Fv fragments, which are capable of binding an epitopic determinant. Antibodies that bind SECP polypeptides can be prepared using intact polypeptides or using fragments containing small peptides of interest as the immunizing antigen. The polypeptide or oligopeptide used to immunize an animal (e.g., a mouse, a rat, or a rabbit) can be derived from the translation of RNA, or synthesized chemically, and can be conjugated to a carrier protein if desired. Commonly used carriers that are chemically coupled to peptides include bovine serum albumin, thyroglobulin, and keyhole limpet hemocyanin (KLH). The coupled peptide is then used to immunize the animal.

The term "antigenic determinant" refers to that region of a molecule (i.e., an epitope) that makes contact with a particular antibody. When a protein or a fragment of a protein is used to immunize a host animal, numerous regions of the protein may induce the production of antibodies which bind specifically to antigenic determinants (particular regions or three-dimensional structures on the protein). An antigenic determinant may compete with the intact antigen (i.e., the immunogen used to elicit the immune response) for binding to an antibody.

The term "antisense" refers to any composition capable of base-pairing with the "sense" (coding) strand of a specific nucleic acid sequence. Antisense compositions may include DNA; RNA; peptide nucleic acid (PNA); oligonucleotides having modified backbone linkages such as phosphorothioates, methylphosphonates, or benzylphosphonates; oligonucleotides having modified

sugar groups such as 2'-methoxyethyl sugars or 2'-methoxyethoxy sugars; or oligonucleotides having modified bases such as 5-methyl cytosine, 2'-deoxyuracil, or 7-deaza-2'-deoxyguanosine. Antisense molecules may be produced by any method including chemical synthesis or transcription. Once introduced into a cell, the complementary antisense molecule base-pairs with a naturally occurring nucleic acid sequence produced by the cell to form duplexes which block either transcription or translation. The designation "negative" or "minus" can refer to the antisense strand, and the designation "positive" or "plus" can refer to the sense strand of a reference DNA molecule.

The term "biologically active" refers to a protein having structural, regulatory, or biochemical functions of a naturally occurring molecule. Likewise, "immunologically active" or "immunogenic" refers to the capability of the natural, recombinant, or synthetic SECP, or of any oligopeptide thereof, to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies.

"Complementary" describes the relationship between two single-stranded nucleic acid sequences that anneal by base-pairing. For example, 5'-AGT-3' pairs with its complement, 3'-TCA-5'.

A "composition comprising a given polynucleotide sequence" and a "composition comprising a given amino acid sequence" refer broadly to any composition containing the given polynucleotide or amino acid sequence. The composition may comprise a dry formulation or an aqueous solution. Compositions comprising polynucleotide sequences encoding SECP or fragments of SECP may be employed as hybridization probes. The probes may be stored in freeze-dried form and may be associated with a stabilizing agent such as a carbohydrate. In hybridizations, the probe may be deployed in an aqueous solution containing salts (e.g., NaCl), detergents (e.g., sodium dodecyl sulfate; SDS), and other components (e.g., Denhardt's solution, dry milk, salmon sperm DNA, etc.).

"Consensus sequence" refers to a nucleic acid sequence which has been subjected to repeated DNA sequence analysis to resolve uncalled bases, extended using the XL-PCR kit (Applied Biosystems, Foster City CA) in the 5' and/or the 3' direction, and resequenced, or which has been assembled from one or more overlapping cDNA, EST, or genomic DNA fragments using a computer program for fragment assembly, such as the GELVIEW fragment assembly system (GCG, Madison WI) or Phrap (University of Washington, Seattle WA). Some sequences have been both extended and assembled to produce the consensus sequence.

"Conservative amino acid substitutions" are those substitutions that are predicted to least interfere with the properties of the original protein, i.e., the structure and especially the function of the protein is conserved and not significantly changed by such substitutions. The table below shows amino acids which may be substituted for an original amino acid in a protein and which are regarded as conservative amino acid substitutions.

Original Residue	Conservative Substitution
Ala	Gly, Ser
Arg	His, Lys
Asn	Asp, Gln, His
Asp	Asn, Glu
Cys	Ala, Ser
Gln	Asn, Glu, His
Glu	Asp, Gln, His
Gly	Ala
His	Asn, Arg, Gln, Glu
Ile	Leu, Val
Leu	Ile, Val
Lys	Arg, Gln, Glu
Met	Leu, Ile
Phe	His, Met, Leu, Trp, Tyr
Ser	Cys, Thr
Thr	Ser, Val
Trp	Phe, Tyr
Tyr	His, Phe, Trp
Val	Ile, Leu, Thr

Conservative amino acid substitutions generally maintain (a) the structure of the polypeptide backbone in the area of the substitution, for example, as a beta sheet or alpha helical conformation, (b) the charge or hydrophobicity of the molecule at the site of the substitution, and/or (c) the bulk of the side chain.

A "deletion" refers to a change in the amino acid or nucleotide sequence that results in the absence of one or more amino acid residues or nucleotides.

The term "derivative" refers to a chemically modified polynucleotide or polypeptide. Chemical modifications of a polynucleotide can include, for example, replacement of hydrogen by an alkyl, acyl, hydroxyl, or amino group. A derivative polynucleotide encodes a polypeptide which retains at least one biological or immunological function of the natural molecule. A derivative polypeptide is one modified by glycosylation, pegylation, or any similar process that retains at least one biological or immunological function of the polypeptide from which it was derived.

A "detectable label" refers to a reporter molecule or enzyme that is capable of generating a measurable signal and is covalently or noncovalently joined to a polynucleotide or polypeptide.

"Differential expression" refers to increased or upregulated; or decreased, downregulated, or absent gene or protein expression, determined by comparing at least two different samples. Such comparisons may be carried out between, for example, a treated and an untreated sample, or a diseased and a normal sample.

A "fragment" is a unique portion of SECP or the polynucleotide encoding SECP which is identical in sequence to but shorter in length than the parent sequence. A fragment may comprise up

to the entire length of the defined sequence, minus one nucleotide/amino acid residue. For example, a fragment may comprise from 5 to 1000 contiguous nucleotides or amino acid residues. A fragment used as a probe, primer, antigen, therapeutic molecule, or for other purposes, may be at least 5, 10, 15, 16, 20, 25, 30, 40, 50, 60, 75, 100, 150, 250 or at least 500 contiguous nucleotides or amino acid residues in length. Fragments may be preferentially selected from certain regions of a molecule. For example, a polypeptide fragment may comprise a certain length of contiguous amino acids selected from the first 250 or 500 amino acids (or first 25% or 50%) of a polypeptide as shown in a certain defined sequence. Clearly these lengths are exemplary, and any length that is supported by the specification, including the Sequence Listing, tables, and figures, may be encompassed by the present embodiments.

A fragment of SEQ ID NO:45-88 comprises a region of unique polynucleotide sequence that specifically identifies SEQ ID NO:45-88, for example, as distinct from any other sequence in the genome from which the fragment was obtained. A fragment of SEQ ID NO:45-88 is useful, for example, in hybridization and amplification technologies and in analogous methods that distinguish SEQ ID NO:45-88 from related polynucleotide sequences. The precise length of a fragment of SEQ ID NO:45-88 and the region of SEQ ID NO:45-88 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

A fragment of SEQ ID NO:1-44 is encoded by a fragment of SEQ ID NO:45-88. A fragment of SEQ ID NO:1-44 comprises a region of unique amino acid sequence that specifically identifies SEQ ID NO:1-44. For example, a fragment of SEQ ID NO:1-44 is useful as an immunogenic peptide for the development of antibodies that specifically recognize SEQ ID NO:1-44. The precise length of a fragment of SEQ ID NO:1-44 and the region of SEQ ID NO:1-44 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

A "full length" polynucleotide sequence is one containing at least a translation initiation codon (e.g., methionine) followed by an open reading frame and a translation termination codon. A "full length" polynucleotide sequence encodes a "full length" polypeptide sequence.

"Homology" refers to sequence similarity or, interchangeably, sequence identity, between two or more polynucleotide sequences or two or more polypeptide sequences.

The terms "percent identity" and "% identity," as applied to polynucleotide sequences, refer to the percentage of residue matches between at least two polynucleotide sequences aligned using a standardized algorithm. Such an algorithm may insert, in a standardized and reproducible way, gaps in the sequences being compared in order to optimize alignment between two sequences, and therefore achieve a more meaningful comparison of the two sequences.

Percent identity between polynucleotide sequences may be determined using the default

parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e sequence alignment program. This program is part of the LASERGENE software package, a suite of molecular biological analysis programs (DNASTAR, Madison WI). CLUSTAL V is described in Higgins, D.G. and P.M. Sharp (1989) CABIOS 5:151-153 and in Higgins, D.G. et al. (1992) CABIOS 8:189-191. For pairwise alignments of polynucleotide sequences, the default parameters are set as follows: Ktuple=2, gap penalty=5, window=4, and "diagonals saved"=4. The "weighted" residue weight table is selected as the default. Percent identity is reported by CLUSTAL V as the "percent similarity" between aligned polynucleotide sequences.

Alternatively, a suite of commonly used and freely available sequence comparison algorithms is provided by the National Center for Biotechnology Information (NCBI) Basic Local Alignment Search Tool (BLAST) (Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410), which is available from several sources, including the NCBI, Bethesda, MD, and on the Internet at http://www.ncbi.nlm.nih.gov/BLAST/. The BLAST software suite includes various sequence analysis programs including "blastn," that is used to align a known polynucleotide sequence with other polynucleotide sequences from a variety of databases. Also available is a tool called "BLAST 2 Sequences" that is used for direct pairwise comparison of two nucleotide sequences. "BLAST 2 Sequences" can be accessed and used interactively at http://www.ncbi.nlm.nih.gov/gorf/bl2.html. The "BLAST 2 Sequences" tool can be used for both blastn and blastp (discussed below). BLAST programs are commonly used with gap and other parameters set to default settings. For example, to compare two nucleotide sequences, one may use blastn with the "BLAST 2 Sequences" tool Version 2.0.12 (April-21-2000) set at default parameters. Such default parameters may be, for example:

Matrix: BLOSUM62
Reward for match: 1
Penalty for mismatch: -2
Open Gap: 5 and Extension Gap: 2 penalties
Gap x drop-off: 50
Expect: 10
Word Size: 11
Filter: on

Percent identity may be measured over the length of an entire defined sequence, for example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for example, over the length of a fragment taken from a larger, defined sequence, for instance, a fragment of at least 20, at least 30, at least 40, at least 50, at least 70, at least 100, or at least 200 contiguous nucleotides. Such lengths are exemplary only, and it is understood that any fragment length supported by the sequences shown herein, in the tables, figures, or Sequence Listing, may be used to

describe a length over which percentage identity may be measured.

Nucleic acid sequences that do not show a high degree of identity may nevertheless encode similar amino acid sequences due to the degeneracy of the genetic code. It is understood that changes in a nucleic acid sequence can be made using this degeneracy to produce multiple nucleic acid sequences that all encode substantially the same protein.

The phrases "percent identity" and "% identity," as applied to polypeptide sequences, refer to the percentage of residue matches between at least two polypeptide sequences aligned using a standardized algorithm. Methods of polypeptide sequence alignment are well-known. Some alignment methods take into account conservative amino acid substitutions. Such conservative substitutions, explained in more detail above, generally preserve the charge and hydrophobicity at the site of substitution, thus preserving the structure (and therefore function) of the polypeptide.

Percent identity between polypeptide sequences may be determined using the default parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e sequence alignment program (described and referenced above). For pairwise alignments of polypeptide sequences using CLUSTAL V, the default parameters are set as follows: Ktuple=1, gap penalty=3, window=5, and "diagonals saved"=5. The PAM250 matrix is selected as the default residue weight table. As with polynucleotide alignments, the percent identity is reported by CLUSTAL V as the "percent similarity" between aligned polypeptide sequence pairs.

Alternatively the NCBI BLAST software suite may be used. For example, for a pairwise comparison of two polypeptide sequences, one may use the "BLAST 2 Sequences" tool Version 2.0.12 (April-21-2000) with blastp set at default parameters. Such default parameters may be, for example:

Matrix: BLOSUM62

Open Gap: 11 and Extension Gap: 1 penalties

Gap x drop-off: 50

Expect: 10

Word Size: 3

Filter: on

Percent identity may be measured over the length of an entire defined polypeptide sequence, for example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for example, over the length of a fragment taken from a larger, defined polypeptide sequence, for instance, a fragment of at least 15, at least 20, at least 30, at least 40, at least 50, at least 70 or at least 150 contiguous residues. Such lengths are exemplary only, and it is understood that any fragment length supported by the sequences shown herein, in the tables, figures or Sequence Listing, may be used to describe a length over which percentage identity may be measured.

"Human artificial chromosomes" (HACs) are linear microchromosomes which may contain DNA sequences of about 6 kb to 10 Mb in size and which contain all of the elements required for chromosome replication, segregation and maintenance.

The term "humanized antibody" refers to an antibody molecule in which the amino acid sequence in the non-antigen binding regions has been altered so that the antibody more closely resembles a human antibody, and still retains its original binding ability.

"Hybridization" refers to the process by which a polynucleotide strand anneals with a complementary strand through base pairing under defined hybridization conditions. Specific hybridization is an indication that two nucleic acid sequences share a high degree of complementarity. Specific hybridization complexes form under permissive annealing conditions and remain hybridized after the "washing" step(s). The washing step(s) is particularly important in determining the stringency of the hybridization process, with more stringent conditions allowing less non-specific binding, i.e., binding between pairs of nucleic acid strands that are not perfectly matched. Permissive conditions for annealing of nucleic acid sequences are routinely determinable by one of ordinary skill in the art and may be consistent among hybridization experiments, whereas wash conditions may be varied among experiments to achieve the desired stringency, and therefore hybridization specificity. Permissive annealing conditions occur, for example, at 68°C in the presence of about 6 x SSC, about 1% (w/v) SDS, and about 100 µg/ml sheared, denatured salmon sperm DNA.

Generally, stringency of hybridization is expressed, in part, with reference to the temperature under which the wash step is carried out. Such wash temperatures are typically selected to be about 5°C to 20°C lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength and pH. The T_m is the temperature (under defined ionic strength and pH) at which 50% of the target sequence hybridizes to a perfectly matched probe. An equation for calculating T_m and conditions for nucleic acid hybridization are well known and can be found in Sambrook, J. et al. (1989) Molecular Cloning: A Laboratory Manual, 2nd ed., vol. 1-3, Cold Spring Harbor Press, Plainview NY; specifically see volume 2, chapter 9.

High stringency conditions for hybridization between polynucleotides of the present invention include wash conditions of 68°C in the presence of about 0.2 x SSC and about 0.1% SDS, for 1 hour. Alternatively, temperatures of about 65°C, 60°C, 55°C, or 42°C may be used. SSC concentration may be varied from about 0.1 to 2 x SSC, with SDS being present at about 0.1%. Typically, blocking reagents are used to block non-specific hybridization. Such blocking reagents include, for instance, sheared and denatured salmon sperm DNA at about 100-200 µg/ml. Organic solvent, such as formamide at a concentration of about 35-50% v/v, may also be used under particular circumstances, such as for RNA:DNA hybridizations. Useful variations on these wash conditions will be readily apparent to those of ordinary skill in the art. Hybridization, particularly under high

stringency conditions, may be suggestive of evolutionary similarity between the nucleotides. Such similarity is strongly indicative of a similar role for the nucleotides and their encoded polypeptides.

The term "hybridization complex" refers to a complex formed between two nucleic acid sequences by virtue of the formation of hydrogen bonds between complementary bases. A hybridization complex may be formed in solution (e.g., C_0t or R_0t analysis) or formed between one nucleic acid sequence present in solution and another nucleic acid sequence immobilized on a solid support (e.g., paper, membranes, filters, chips, pins or glass slides, or any other appropriate substrate to which cells or their nucleic acids have been fixed).

The words "insertion" and "addition" refer to changes in an amino acid or nucleotide sequence resulting in the addition of one or more amino acid residues or nucleotides, respectively.

"Immune response" can refer to conditions associated with inflammation, trauma, immune disorders, or infectious or genetic disease, etc. These conditions can be characterized by expression of various factors, e.g., cytokines, chemokines, and other signaling molecules, which may affect cellular and systemic defense systems.

An "immunogenic fragment" is a polypeptide or oligopeptide fragment of SECP which is capable of eliciting an immune response when introduced into a living organism, for example, a mammal. The term "immunogenic fragment" also includes any polypeptide or oligopeptide fragment of SECP which is useful in any of the antibody production methods disclosed herein or known in the art.

The term "microarray" refers to an arrangement of a plurality of polynucleotides, polypeptides, or other chemical compounds on a substrate.

The terms "element" and "array element" refer to a polynucleotide, polypeptide, or other chemical compound having a unique and defined position on a microarray.

The term "modulate" refers to a change in the activity of SECP. For example, modulation may cause an increase or a decrease in protein activity, binding characteristics, or any other biological, functional, or immunological properties of SECP.

The phrases "nucleic acid" and "nucleic acid sequence" refer to a nucleotide, oligonucleotide, polynucleotide, or any fragment thereof. These phrases also refer to DNA or RNA of genomic or synthetic origin which may be single-stranded or double-stranded and may represent the sense or the antisense strand, to peptide nucleic acid (PNA), or to any DNA-like or RNA-like material.

"Operably linked" refers to the situation in which a first nucleic acid sequence is placed in a functional relationship with a second nucleic acid sequence. For instance, a promoter is operably linked to a coding sequence if the promoter affects the transcription or expression of the coding sequence. Operably linked DNA sequences may be in close proximity or contiguous and, where necessary to join two protein coding regions, in the same reading frame.

"Peptide nucleic acid" (PNA) refers to an antisense molecule or anti-gene agent which comprises an oligonucleotide of at least about 5 nucleotides in length linked to a peptide backbone of amino acid residues ending in lysine. The terminal lysine confers solubility to the composition. PNAs preferentially bind complementary single stranded DNA or RNA and stop transcript elongation, and may be pegylated to extend their lifespan in the cell.

"Post-translational modification" of an SECP may involve lipidation, glycosylation, phosphorylation, acetylation, racemization, proteolytic cleavage, and other modifications known in the art. These processes may occur synthetically or biochemically. Biochemical modifications will vary by cell type depending on the enzymatic milieu of SECP.

"Probe" refers to nucleic acid sequences encoding SECP, their complements, or fragments thereof, which are used to detect identical, allelic or related nucleic acid sequences. Probes are isolated oligonucleotides or polynucleotides attached to a detectable label or reporter molecule. Typical labels include radioactive isotopes, ligands, chemiluminescent agents, and enzymes. "Primers" are short nucleic acids, usually DNA oligonucleotides, which may be annealed to a target polynucleotide by complementary base-pairing. The primer may then be extended along the target DNA strand by a DNA polymerase enzyme. Primer pairs can be used for amplification (and identification) of a nucleic acid sequence, e.g., by the polymerase chain reaction (PCR).

Probes and primers as used in the present invention typically comprise at least 15 contiguous nucleotides of a known sequence. In order to enhance specificity, longer probes and primers may also be employed, such as probes and primers that comprise at least 20, 25, 30, 40, 50, 60, 70, 80, 90, 100, or at least 150 consecutive nucleotides of the disclosed nucleic acid sequences. Probes and primers may be considerably longer than these examples, and it is understood that any length supported by the specification, including the tables, figures, and Sequence Listing, may be used.

Methods for preparing and using probes and primers are described in the references, for example Sambrook, J. et al. (1989) Molecular Cloning: A Laboratory Manual, 2nd ed., vol. 1-3, Cold Spring Harbor Press, Plainview NY; Ausubel, F.M. et al. (1987) Current Protocols in Molecular Biology, Greene Publ. Assoc. & Wiley-Intersciences, New York NY; Innis, M. et al. (1990) PCR Protocols, A Guide to Methods and Applications, Academic Press, San Diego CA. PCR primer pairs can be derived from a known sequence, for example, by using computer programs intended for that purpose such as Primer (Version 0.5, 1991, Whitehead Institute for Biomedical Research, Cambridge MA).

Oligonucleotides for use as primers are selected using software known in the art for such purpose. For example, OLIGO 4.06 software is useful for the selection of PCR primer pairs of up to 100 nucleotides each, and for the analysis of oligonucleotides and larger polynucleotides of up to 5,000 nucleotides from an input polynucleotide sequence of up to 32 kilobases. Similar primer

selection programs have incorporated additional features for expanded capabilities. For example, the PrimOU primer selection program (available to the public from the Genome Center at University of Texas South West Medical Center, Dallas TX) is capable of choosing specific primers from megabase sequences and is thus useful for designing primers on a genome-wide scope. The Primer3 primer selection program (available to the public from the Whitehead Institute/MIT Center for Genome Research, Cambridge MA) allows the user to input a "mispriming library," in which sequences to avoid as primer binding sites are user-specified. Primer3 is useful, in particular, for the selection of oligonucleotides for microarrays. (The source code for the latter two primer selection programs may also be obtained from their respective sources and modified to meet the user's specific needs.) The PrimeGen program (available to the public from the UK Human Genome Mapping Project Resource Centre, Cambridge UK) designs primers based on multiple sequence alignments, thereby allowing selection of primers that hybridize to either the most conserved or least conserved regions of aligned nucleic acid sequences. Hence, this program is useful for identification of both unique and conserved oligonucleotides and polynucleotide fragments. The oligonucleotides and polynucleotide fragments identified by any of the above selection methods are useful in hybridization technologies, for example, as PCR or sequencing primers, microarray elements, or specific probes to identify fully or partially complementary polynucleotides in a sample of nucleic acids. Methods of oligonucleotide selection are not limited to those described above.

A "recombinant nucleic acid" is a sequence that is not naturally occurring or has a sequence that is made by an artificial combination of two or more otherwise separated segments of sequence. This artificial combination is often accomplished by chemical synthesis or, more commonly, by the artificial manipulation of isolated segments of nucleic acids, e.g., by genetic engineering techniques such as those described in Sambrook, supra. The term recombinant includes nucleic acids that have been altered solely by addition, substitution, or deletion of a portion of the nucleic acid. Frequently, a recombinant nucleic acid may include a nucleic acid sequence operably linked to a promoter sequence. Such a recombinant nucleic acid may be part of a vector that is used, for example, to transform a cell.

Alternatively, such recombinant nucleic acids may be part of a viral vector, e.g., based on a vaccinia virus, that could be use to vaccinate a mammal wherein the recombinant nucleic acid is expressed, inducing a protective immunological response in the mammal.

A "regulatory element" refers to a nucleic acid sequence usually derived from untranslated regions of a gene and includes enhancers, promoters, introns, and 5' and 3' untranslated regions (UTRs). Regulatory elements interact with host or viral proteins which control transcription, translation, or RNA stability.

"Reporter molecules" are chemical or biochemical moieties used for labeling a nucleic acid,

amino acid, or antibody. Reporter molecules include radionuclides; enzymes; fluorescent, chemiluminescent, or chromogenic agents; substrates; cofactors; inhibitors; magnetic particles; and other moieties known in the art.

An "RNA equivalent," in reference to a DNA sequence, is composed of the same linear sequence of nucleotides as the reference DNA sequence with the exception that all occurrences of the nitrogenous base thymine are replaced with uracil, and the sugar backbone is composed of ribose instead of deoxyribose.

The term "sample" is used in its broadest sense. A sample suspected of containing SECP, nucleic acids encoding SECP, or fragments thereof may comprise a bodily fluid; an extract from a cell, chromosome, organelle, or membrane isolated from a cell; a cell; genomic DNA, RNA, or cDNA, in solution or bound to a substrate; a tissue; a tissue print; etc.

The terms "specific binding" and "specifically binding" refer to that interaction between a protein or peptide and an agonist, an antibody, an antagonist, a small molecule, or any natural or synthetic binding composition. The interaction is dependent upon the presence of a particular structure of the protein, e.g., the antigenic determinant or epitope, recognized by the binding molecule. For example, if an antibody is specific for epitope "A," the presence of a polypeptide comprising the epitope A, or the presence of free unlabeled A, in a reaction containing free labeled A and the antibody will reduce the amount of labeled A that binds to the antibody.

The term "substantially purified" refers to nucleic acid or amino acid sequences that are removed from their natural environment and are isolated or separated, and are at least 60% free, preferably at least 75% free, and most preferably at least 90% free from other components with which they are naturally associated.

A "substitution" refers to the replacement of one or more amino acid residues or nucleotides by different amino acid residues or nucleotides, respectively.

"Substrate" refers to any suitable rigid or semi-rigid support including membranes, filters, chips, slides, wafers, fibers, magnetic or nonmagnetic beads, gels, tubing, plates, polymers, microparticles and capillaries. The substrate can have a variety of surface forms, such as wells, trenches, pins, channels and pores, to which polynucleotides or polypeptides are bound.

A "transcript image" refers to the collective pattern of gene expression by a particular cell type or tissue under given conditions at a given time.

"Transformation" describes a process by which exogenous DNA is introduced into a recipient cell. Transformation may occur under natural or artificial conditions according to various methods well known in the art, and may rely on any known method for the insertion of foreign nucleic acid sequences into a prokaryotic or eukaryotic host cell. The method for transformation is selected based on the type of host cell being transformed and may include, but is not limited to, bacteriophage or

viral infection, electroporation, heat shock, lipofection, and particle bombardment. The term "transformed cells" includes stably transformed cells in which the inserted DNA is capable of replication either as an autonomously replicating plasmid or as part of the host chromosome, as well as transiently transformed cells which express the inserted DNA or RNA for limited periods of time.

A "transgenic organism," as used herein, is any organism, including but not limited to animals and plants, in which one or more of the cells of the organism contains heterologous nucleic acid introduced by way of human intervention, such as by transgenic techniques well known in the art. The nucleic acid is introduced into the cell, directly or indirectly by introduction into a precursor of the cell, by way of deliberate genetic manipulation, such as by microinjection or by infection with a recombinant virus. The term genetic manipulation does not include classical cross-breeding, or in vitro fertilization, but rather is directed to the introduction of a recombinant DNA molecule. The transgenic organisms contemplated in accordance with the present invention include bacteria, cyanobacteria, fungi, plants and animals. The isolated DNA of the present invention can be introduced into the host by methods known in the art, for example infection, transfection, transformation or transconjugation. Techniques for transferring the DNA of the present invention into such organisms are widely known and provided in references such as Sambrook et al. (1989), supra.

A "variant" of a particular nucleic acid sequence is defined as a nucleic acid sequence having at least 40% sequence identity to the particular nucleic acid sequence over a certain length of one of the nucleic acid sequences using blastn with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such a pair of nucleic acids may show, for example, at least 50%, at least 60%, at least 70%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% or greater sequence identity over a certain defined length. A variant may be described as, for example, an "allelic" (as defined above), "splice," "species," or "polymorphic" variant. A splice variant may have significant identity to a reference molecule, but will generally have a greater or lesser number of polynucleotides due to alternative splicing of exons during mRNA processing. The corresponding polypeptide may possess additional functional domains or lack domains that are present in the reference molecule. Species variants are polynucleotide sequences that vary from one species to another. The resulting polypeptides will generally have significant amino acid identity relative to each other. A polymorphic variant is a variation in the polynucleotide sequence of a particular gene between individuals of a given species. Polymorphic variants also may encompass "single nucleotide polymorphisms" (SNPs) in which the polynucleotide sequence varies by one nucleotide base. The presence of SNPs may be indicative of, for example, a certain population, a disease state, or a propensity for a disease state.

A "variant" of a particular polypeptide sequence is defined as a polypeptide sequence having at least 40% sequence identity to the particular polypeptide sequence over a certain length of one of the polypeptide sequences using blastp with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such a pair of polypeptides may show, for example, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% or greater sequence identity over a certain defined length of one of the polypeptides.

THE INVENTION

The invention is based on the discovery of new human secreted proteins (SECP), the polynucleotides encoding SECP, and the use of these compositions for the diagnosis, treatment, or prevention of cell proliferative, autoimmune/inflammatory, cardiovascular, neurological, and developmental disorders.

Table 1 summarizes the nomenclature for the full length polynucleotide and polypeptide sequences of the invention. Each polynucleotide and its corresponding polypeptide are correlated to a single Incyte project identification number (Incyte Project ID). Each polypeptide sequence is denoted by both a polypeptide sequence identification number (Polypeptide SEQ ID NO:) and an Incyte polypeptide sequence number (Incyte Polypeptide ID) as shown. Each polynucleotide sequence is denoted by both a polynucleotide sequence identification number (Polynucleotide SEQ ID NO:) and an Incyte polynucleotide consensus sequence number (Incyte Polynucleotide ID) as shown.

Table 2 shows sequences with homology to the polypeptides of the invention as identified by BLAST analysis against the GenBank protein (genpept) database. Columns 1 and 2 show the polypeptide sequence identification number (Polypeptide SEQ ID NO:) and the corresponding Incyte polypeptide sequence number (Incyte Polypeptide ID) for polypeptides of the invention. Column 3 shows the GenBank identification number (Genbank ID NO:) of the nearest GenBank homolog. Column 4 shows the probability score for the match between each polypeptide and its GenBank homolog. Column 5 shows the annotation of the GenBank homolog along with relevant citations where applicable, all of which are expressly incorporated by reference herein.

Table 3 shows various structural features of the polypeptides of the invention. Columns 1 and 2 show the polypeptide sequence identification number (SEQ ID NO:) and the corresponding Incyte polypeptide sequence number (Incyte Polypeptide ID) for each polypeptide of the invention. Column 3 shows the number of amino acid residues in each polypeptide. Column 4 shows potential phosphorylation sites, and column 5 shows potential glycosylation sites, as determined by the MOTIFS program of the GCG sequence analysis software package (Genetics Computer Group, Madison WI). Column 6 shows amino acid residues comprising signature sequences, domains, and

motifs. In particular, the locations of signal peptides (as indicated by "Signal_peptide" or "Signal_cleavage") are shown. Column 7 shows analytical methods for protein structure/function analysis and in some cases, searchable databases to which the analytical methods were applied.

Together, tables 2 and 3 summarize the properties of each polypeptide of the invention, and these properties establish that the claimed polypeptides are secreted proteins. For example, SEQ ID NO:1 is 51% identical to human UDP-GalNAc:polypeptide N-acetylgalactosaminyltransferase (GenBank ID g971461) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is 1.5e-141, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:1 also contains a signal peptide and a transmembrane domain as determined by hidden Markov model (HMM)-based methods. (See Table 3.) Likewise, SPScan analysis also indicates the presence of an N-terminal signal peptide in SEQ ID NO:1. Taken together, the evidence shows that SEQ ID NO:1 is present in the secretory pathway as an N-acetylgalactosaminyl transferase.

For example, SEQ ID NO:2 is 90% identical to mouse seizure-related gene product 6 type 2 precursor (GenBank ID g1139548) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is 0.0, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:2 also contains five sushi domains and two CUB domains as determined by searching for statistically significant matches in the hidden Markov model (HMM)-based PFAM database of conserved protein family domains. (See Table 3.) In addition, SEQ ID NO:2 contains a signal peptide and a single transmembrane domain, as identified by HMMER analysis.

For example, SEQ ID NO:3 is 43% identical to Gallus gallus lysozyme (GenBank ID g4467410) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is 5.2e-40, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:3 also contains a G-lysozyme signature domain as determined by searching for statistically significant matches in the BLIMPS analysis of the PRINTS database of conserved protein motifs. (See Table 3.) Data from the PFAM, PRODOM and DOMO databases provide further corroborative evidence that SEQ ID NO:3 is a lysozyme.

For example, SEQ ID NO:17 has a signal peptide, as determined by SPScan and hidden Markov model (HMM) based analyses. SEQ ID NO:17 is 86% identical to human immunoglobulin lambda light chain (GenBank ID g33702) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is 2.2e-106, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:17 also contains an immunoglobulin domain as determined by searching for statistically significant matches in the HMM-based PFAM database of conserved protein family domains. (See Table 3.) Data from

BLIMPS, MOTIFS, and PROFILESCAN analyses provide further corroborative evidence that SEQ ID NO:17 is a secreted immunoglobulin. The available evidence shows that SEQ ID NO:19 is also a secreted immunoglobulin.

For example, SEQ ID NO:38 shows 95% identity to human immunoglobulin lambda light chain (GenBank ID g33718) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is 5.2e-114, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:38 also contains an immunoglobulin domain as determined by searching for statistically significant matches in the hidden Markov model (HMM)-based PFAM database of conserved protein family domains. (See Table 3.) Data from BLIMPS, MOTIFS, and PROFILESCAN analyses provide further corroborative evidence that SEQ ID NO:38 is a secreted protein, and more specifically an immunoglobulin. SEQ ID NO:4-16, SEQ ID NO:18-37, and SEQ ID NO:39-44 were analyzed and annotated in a similar manner. The algorithms and parameters for the analysis of SEQ ID NO:1-44 are described in Table 7.

As shown in Table 4, the full length polynucleotide sequences of the present invention were assembled using cDNA sequences or coding (exon) sequences derived from genomic DNA, or any combination of these two types of sequences. Columns 1 and 2 list the polynucleotide sequence identification number (Polynucleotide SEQ ID NO:) and the corresponding Incyte polynucleotide consensus sequence number (Incyte Polynucleotide ID) for each polynucleotide of the invention. Column 3 shows the length of each polynucleotide sequence in basepairs. Column 4 lists fragments of the polynucleotide sequences which are useful, for example, in hybridization or amplification technologies that identify SEQ ID NO:45-88 or that distinguish between SEQ ID NO:45-88 and related polynucleotide sequences. Column 5 shows identification numbers corresponding to cDNA sequences, coding sequences (exons) predicted from genomic DNA, and/or sequence assemblages comprised of both cDNA and genomic DNA. These sequences were used to assemble the full length polynucleotide sequences of the invention. Columns 6 and 7 of Table 4 show the nucleotide start (5′) and stop (3′) positions of the cDNA and/or genomic sequences in column 5 relative to their respective full length sequences.

The identification numbers in Column 5 of Table 4 may refer specifically, for example, to Incyte cDNAs along with their corresponding cDNA libraries. For example, 6735891H1 is the identification number of an Incyte cDNA sequence, and LIVRTUT13 is the cDNA library from which it is derived. Incyte cDNAs for which cDNA libraries are not indicated were derived from pooled cDNA libraries (e.g., 71013085V1). Alternatively, the identification numbers in column 5 may refer to GenBank cDNAs or ESTs (e.g., g1496797) which contributed to the assembly of the full length polynucleotide sequences. Alternatively, the identification numbers in column 5 may refer to coding regions predicted by Genscan analysis of genomic DNA. The Genscan-predicted coding sequences

may have been edited prior to assembly. (See Example IV.) Alternatively, the identification numbers in column 5 may refer to assemblages of both cDNA and Genscan-predicted exons brought together by an "exon stitching" algorithm. (See Example V.) Alternatively, the identification numbers in column 5 may refer to assemblages of both cDNA and Genscan-predicted exons brought together by an "exon-stretching" algorithm. (See Example V.) In some cases, Incyte cDNA coverage redundant with the sequence coverage shown in column 5 was obtained to confirm the final consensus polynucleotide sequence, but the relevant Incyte cDNA identification numbers are not shown.

Table 5 shows the representative cDNA libraries for those full length polynucleotide sequences which were assembled using Incyte cDNA sequences. The representative cDNA library is the Incyte cDNA library which is most frequently represented by the Incyte cDNA sequences which were used to assemble and confirm the above polynucleotide sequences. The tissues and vectors which were used to construct the cDNA libraries shown in Table 5 are described in Table 6.

The invention also encompasses SECP variants. A preferred SECP variant is one which has at least about 80%, or alternatively at least about 90%, or even at least about 95% amino acid sequence identity to the SECP amino acid sequence, and which contains at least one functional or structural characteristic of SECP.

The invention also encompasses polynucleotides which encode SECP. In a particular embodiment, the invention encompasses a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:45-88, which encodes SECP. The polynucleotide sequences of SEQ ID NO:45-88, as presented in the Sequence Listing, embrace the equivalent RNA sequences, wherein occurrences of the nitrogenous base thymine are replaced with uracil, and the sugar backbone is composed of ribose instead of deoxyribose.

The invention also encompasses a variant of a polynucleotide sequence encoding SECP. In particular, such a variant polynucleotide sequence will have at least about 70%, or alternatively at least about 85%, or even at least about 95% polynucleotide sequence identity to the polynucleotide sequence encoding SECP. A particular aspect of the invention encompasses a variant of a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:45-88 which has at least about 70%, or alternatively at least about 85%, or even at least about 95% polynucleotide sequence identity to a nucleic acid sequence selected from the group consisting of SEQ ID NO:45-88. Any one of the polynucleotide variants described above can encode an amino acid sequence which contains at least one functional or structural characteristic of SECP.

It will be appreciated by those skilled in the art that as a result of the degeneracy of the genetic code, a multitude of polynucleotide sequences encoding SECP, some bearing minimal similarity to the polynucleotide sequences of any known and naturally occurring gene, may be produced. Thus, the invention contemplates each and every possible variation of polynucleotide

sequence that could be made by selecting combinations based on possible codon choices. These combinations are made in accordance with the standard triplet genetic code as applied to the polynucleotide sequence of naturally occurring SECP, and all such variations are to be considered as being specifically disclosed.

Although nucleotide sequences which encode SECP and its variants are generally capable of hybridizing to the nucleotide sequence of the naturally occurring SECP under appropriately selected conditions of stringency, it may be advantageous to produce nucleotide sequences encoding SECP or its derivatives possessing a substantially different codon usage, e.g., inclusion of non-naturally occurring codons. Codons may be selected to increase the rate at which expression of the peptide occurs in a particular prokaryotic or eukaryotic host in accordance with the frequency with which particular codons are utilized by the host. Other reasons for substantially altering the nucleotide sequence encoding SECP and its derivatives without altering the encoded amino acid sequences include the production of RNA transcripts having more desirable properties, such as a greater half-life, than transcripts produced from the naturally occurring sequence.

The invention also encompasses production of DNA sequences which encode SECP and SECP derivatives, or fragments thereof, entirely by synthetic chemistry. After production, the synthetic sequence may be inserted into any of the many available expression vectors and cell systems using reagents well known in the art. Moreover, synthetic chemistry may be used to introduce mutations into a sequence encoding SECP or any fragment thereof.

Also encompassed by the invention are polynucleotide sequences that are capable of hybridizing to the claimed polynucleotide sequences, and, in particular, to those shown in SEQ ID NO:45-88 and fragments thereof under various conditions of stringency. (See, e.g., Wahl, G.M. and S.L. Berger (1987) Methods Enzymol. 152:399-407; Kimmel, A.R. (1987) Methods Enzymol. 152:507-511.) Hybridization conditions, including annealing and wash conditions, are described in "Definitions."

Methods for DNA sequencing are well known in the art and may be used to practice any of the embodiments of the invention. The methods may employ such enzymes as the Klenow fragment of DNA polymerase I, SEQUENASE (US Biochemical, Cleveland OH), Taq polymerase (Applied Biosystems), thermostable T7 polymerase (Amersham Pharmacia Biotech, Piscataway NJ), or combinations of polymerases and proofreading exonucleases such as those found in the ELONGASE amplification system (Life Technologies, Gaithersburg MD). Preferably, sequence preparation is automated with machines such as the MICROLAB 2200 liquid transfer system (Hamilton, Reno NV), PTC200 thermal cycler (MJ Research, Watertown MA) and ABI CATALYST 800 thermal cycler (Applied Biosystems). Sequencing is then carried out using either the ABI 373 or 377 DNA sequencing system (Applied Biosystems), the MEGABACE 1000 DNA sequencing system

(Molecular Dynamics, Sunnyvale CA), or other systems known in the art. The resulting sequences are analyzed using a variety of algorithms which are well known in the art. (See, e.g., Ausubel, F.M. (1997) Short Protocols in Molecular Biology, John Wiley & Sons, New York NY, unit 7.7; Meyers, R.A. (1995) Molecular Biology and Biotechnology, Wiley VCH, New York NY, pp. 856-853.)

The nucleic acid sequences encoding SECP may be extended utilizing a partial nucleotide sequence and employing various PCR-based methods known in the art to detect upstream sequences, such as promoters and regulatory elements. For example, one method which may be employed, restriction-site PCR, uses universal and nested primers to amplify unknown sequence from genomic DNA within a cloning vector. (See, e.g., Sarkar, G. (1993) PCR Methods Applic. 2:318-322.) Another method, inverse PCR, uses primers that extend in divergent directions to amplify unknown sequence from a circularized template. The template is derived from restriction fragments comprising a known genomic locus and surrounding sequences. (See, e.g., Triglia, T. et al. (1988) Nucleic Acids Res. 16:8186.) A third method, capture PCR, involves PCR amplification of DNA fragments adjacent to known sequences in human and yeast artificial chromosome DNA. (See, e.g., Lagerstrom, M. et al. (1991) PCR Methods Applic. 1:111-119.) In this method, multiple restriction enzyme digestions and ligations may be used to insert an engineered double-stranded sequence into a region of unknown sequence before performing PCR. Other methods which may be used to retrieve unknown sequences are known in the art. (See, e.g., Parker, J.D. et al. (1991) Nucleic Acids Res. 19:3055-3060). Additionally, one may use PCR, nested primers, and PROMOTERFINDER libraries (Clontech, Palo Alto CA) to walk genomic DNA. This procedure avoids the need to screen libraries and is useful in finding intron/exon junctions. For all PCR-based methods, primers may be designed using commercially available software, such as OLIGO 4.06 primer analysis software (National Biosciences, Plymouth MN) or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the template at temperatures of about 68°C to 72°C.

When screening for full length cDNAs, it is preferable to use libraries that have been size-selected to include larger cDNAs. In addition, random-primed libraries, which often include sequences containing the 5' regions of genes, are preferable for situations in which an oligo d(T) library does not yield a full-length cDNA. Genomic libraries may be useful for extension of sequence into 5' non-transcribed regulatory regions.

Capillary electrophoresis systems which are commercially available may be used to analyze the size or confirm the nucleotide sequence of sequencing or PCR products. In particular, capillary sequencing may employ flowable polymers for electrophoretic separation, four different nucleotide-specific, laser-stimulated fluorescent dyes, and a charge coupled device camera for detection of the emitted wavelengths. Output/light intensity may be converted to electrical signal using appropriate

software (e.g., GENOTYPER and SEQUENCE NAVIGATOR, Applied Biosystems), and the entire process from loading of samples to computer analysis and electronic data display may be computer controlled. Capillary electrophoresis is especially preferable for sequencing small DNA fragments which may be present in limited amounts in a particular sample.

In another embodiment of the invention, polynucleotide sequences or fragments thereof which encode SECP may be cloned in recombinant DNA molecules that direct expression of SECP, or fragments or functional equivalents thereof, in appropriate host cells. Due to the inherent degeneracy of the genetic code, other DNA sequences which encode substantially the same or a functionally equivalent amino acid sequence may be produced and used to express SECP.

The nucleotide sequences of the present invention can be engineered using methods generally known in the art in order to alter SECP-encoding sequences for a variety of purposes including, but not limited to, modification of the cloning, processing, and/or expression of the gene product. DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides may be used to engineer the nucleotide sequences. For example, oligonucleotide-mediated site-directed mutagenesis may be used to introduce mutations that create new restriction sites, alter glycosylation patterns, change codon preference, produce splice variants, and so forth.

The nucleotides of the present invention may be subjected to DNA shuffling techniques such as MOLECULARBREEDING (Maxygen Inc., Santa Clara CA; described in U.S. Patent Number 5,837,458; Chang, C.-C. et al. (1999) Nat. Biotechnol. 17:793-797; Christians, F.C. et al. (1999) Nat. Biotechnol. 17:259-264; and Crameri, A. et al. (1996) Nat. Biotechnol. 14:315-319) to alter or improve the biological properties of SECP, such as its biological or enzymatic activity or its ability to bind to other molecules or compounds. DNA shuffling is a process by which a library of gene variants is produced using PCR-mediated recombination of gene fragments. The library is then subjected to selection or screening procedures that identify those gene variants with the desired properties. These preferred variants may then be pooled and further subjected to recursive rounds of DNA shuffling and selection/screening. Thus, genetic diversity is created through "artificial" breeding and rapid molecular evolution. For example, fragments of a single gene containing random point mutations may be recombined, screened, and then reshuffled until the desired properties are optimized. Alternatively, fragments of a given gene may be recombined with fragments of homologous genes in the same gene family, either from the same or different species, thereby maximizing the genetic diversity of multiple naturally occurring genes in a directed and controllable manner.

In another embodiment, sequences encoding SECP may be synthesized, in whole or in part, using chemical methods well known in the art. (See, e.g., Caruthers, M.H. et al. (1980) Nucleic Acids Symp. Ser. 7:215-223; and Horn, T. et al. (1980) Nucleic Acids Symp. Ser. 7:225-232.)

Alternatively, SECP itself or a fragment thereof may be synthesized using chemical methods. For example, peptide synthesis can be performed using various solution-phase or solid-phase techniques. (See, e.g., Creighton, T. (1984) Proteins, Structures and Molecular Properties, WH Freeman, New York NY, pp. 55-60; and Roberge, J.Y. et al. (1995) Science 269:202-204.) Automated synthesis may be achieved using the ABI 431A peptide synthesizer (Applied Biosystems). Additionally, the amino acid sequence of SECP, or any part thereof, may be altered during direct synthesis and/or combined with sequences from other proteins, or any part thereof, to produce a variant polypeptide or a polypeptide having a sequence of a naturally occurring polypeptide.

The peptide may be substantially purified by preparative high performance liquid chromatography. (See, e.g., Chiez, R.M. and F.Z. Regnier (1990) Methods Enzymol. 182:392-421.) The composition of the synthetic peptides may be confirmed by amino acid analysis or by sequencing. (See, e.g., Creighton, supra, pp. 28-53.)

In order to express a biologically active SECP, the nucleotide sequences encoding SECP or derivatives thereof may be inserted into an appropriate expression vector, i.e., a vector which contains the necessary elements for transcriptional and translational control of the inserted coding sequence in a suitable host. These elements include regulatory sequences, such as enhancers, constitutive and inducible promoters, and 5' and 3' untranslated regions in the vector and in polynucleotide sequences encoding SECP. Such elements may vary in their strength and specificity. Specific initiation signals may also be used to achieve more efficient translation of sequences encoding SECP. Such signals include the ATG initiation codon and adjacent sequences, e.g. the Kozak sequence. In cases where sequences encoding SECP and its initiation codon and upstream regulatory sequences are inserted into the appropriate expression vector, no additional transcriptional or translational control signals may be needed. However, in cases where only coding sequence, or a fragment thereof, is inserted, exogenous translational control signals including an in-frame ATG initiation codon should be provided by the vector. Exogenous translational elements and initiation codons may be of various origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of enhancers appropriate for the particular host cell system used. (See, e.g., Scharf, D. et al. (1994) Results Probl. Cell Differ. 20:125-162.)

Methods which are well known to those skilled in the art may be used to construct expression vectors containing sequences encoding SECP and appropriate transcriptional and translational control elements. These methods include in vitro recombinant DNA techniques, synthetic techniques, and in vivo genetic recombination. (See, e.g., Sambrook, J. et al. (1989) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Plainview NY, ch. 4, 8, and 16-17; Ausubel, F.M. et al. (1995) Current Protocols in Molecular Biology, John Wiley & Sons, New York NY, ch. 9, 13, and 16.)

A variety of expression vector/host systems may be utilized to contain and express sequences

encoding SECP. These include, but are not limited to, microorganisms such as bacteria transformed with recombinant bacteriophage, plasmid; or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with viral expression vectors (e.g., baculovirus); plant cell systems transformed with viral expression vectors (e.g., cauliflower mosaic virus, CaMV, or tobacco mosaic virus, TMV) or with bacterial expression vectors (e.g., Ti or pBR322 plasmids); or animal cell systems. (See, e.g., Sambrook, supra; Ausubel, supra; Van Heeke, G. and S.M. Schuster (1989) J. Biol. Chem. 264:5503-5509; Engelhard, E.K. et al. (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. et al. (1996) Hum. Gene Ther. 7:1937-1945; Takamatsu, N. (1987) EMBO J. 6:307-311; The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York NY, pp. 191-196; Logan, J. and T. Shenk (1984) Proc. Natl. Acad. Sci. USA 81:3655-3659; and Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355.) Expression vectors derived from retroviruses, adenoviruses, or herpes or vaccinia viruses, or from various bacterial plasmids, may be used for delivery of nucleotide sequences to the targeted organ, tissue, or cell population. (See, e.g., Di Nicola, M. et al. (1998) Cancer Gen. Ther. 5(6):350-356; Yu, M. et al. (1993) Proc. Natl. Acad. Sci. USA 90(13):6340-6344; Buller, R.M. et al. (1985) Nature 317(6040):813-815; McGregor, D.P. et al. (1994) Mol. Immunol. 31(3):219-226; and Verma, I.M. and N. Somia (1997) Nature 389:239-242.) The invention is not limited by the host cell employed.

In bacterial systems, a number of cloning and expression vectors may be selected depending upon the use intended for polynucleotide sequences encoding SECP. For example, routine cloning, subcloning, and propagation of polynucleotide sequences encoding SECP can be achieved using a multifunctional E. coli vector such as PBLUESCRIPT (Stratagene, La Jolla CA) or PSPORT1 plasmid (Life Technologies). Ligation of sequences encoding SECP into the vector's multiple cloning site disrupts the *lacZ* gene, allowing a colorimetric screening procedure for identification of transformed bacteria containing recombinant molecules. In addition, these vectors may be useful for in vitro transcription, dideoxy sequencing, single strand rescue with helper phage, and creation of nested deletions in the cloned sequence. (See, e.g., Van Heeke, G. and S.M. Schuster (1989) J. Biol. Chem. 264:5503-5509.) When large quantities of SECP are needed, e.g. for the production of antibodies, vectors which direct high level expression of SECP may be used. For example, vectors containing the strong, inducible SP6 or T7 bacteriophage promoter may be used.

Yeast expression systems may be used for production of SECP. A number of vectors containing constitutive or inducible promoters, such as alpha factor, alcohol oxidase, and PGH promoters, may be used in the yeast Saccharomyces cerevisiae or Pichia pastoris. In addition, such vectors direct either the secretion or intracellular retention of expressed proteins and enable integration of foreign sequences into the host genome for stable propagation. (See, e.g., Ausubel, 1995, supra; Bitter, G.A. et al. (1987) Methods Enzymol. 153:516-544; and Scorer, C.A. et al. (1994)

Bio/Technology 12:181-184.)

Plant systems may also be used for expression of SECP. Transcription of sequences encoding SECP may be driven by viral promoters, e.g., the 35S and 19S promoters of CaMV used alone or in combination with the omega leader sequence from TMV (Takamatsu, N. (1987) EMBO J. 6:307-311). Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters may be used. (See, e.g., Coruzzi, G. et al. (1984) EMBO J. 3:1671-1680; Broglie, R. et al. (1984) Science 224:838-843; and Winter, J. et al. (1991) Results Probl. Cell Differ. 17:85-105.) These constructs can be introduced into plant cells by direct DNA transformation or pathogen-mediated transfection. (See, e.g., The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York NY, pp. 191-196.)

In mammalian cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, sequences encoding SECP may be ligated into an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader sequence. Insertion in a non-essential E1 or E3 region of the viral genome may be used to obtain infective virus which expresses SECP in host cells. (See, e.g., Logan, J. and T. Shenk (1984) Proc. Natl. Acad. Sci. USA 81:3655-3659.) In addition, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, may be used to increase expression in mammalian host cells. SV40 or EBV-based vectors may also be used for high-level protein expression.

Human artificial chromosomes (HACs) may also be employed to deliver larger fragments of DNA than can be contained in and expressed from a plasmid. HACs of about 6 kb to 10 Mb are constructed and delivered via conventional delivery methods (liposomes, polycationic amino polymers, or vesicles) for therapeutic purposes. (See, e.g., Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355.)

For long term production of recombinant proteins in mammalian systems, stable expression of SECP in cell lines is preferred. For example, sequences encoding SECP can be transformed into cell lines using expression vectors which may contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells may be allowed to grow for about 1 to 2 days in enriched media before being switched to selective media. The purpose of the selectable marker is to confer resistance to a selective agent, and its presence allows growth and recovery of cells which successfully express the introduced sequences. Resistant clones of stably transformed cells may be propagated using tissue culture techniques appropriate to the cell type.

Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase and adenine phosphoribosyltransferase genes, for use in *tk*$^-$ and *apr*$^-$ cells, respectively. (See, e.g., Wigler, M. et

al. (1977) Cell 11:223-232; Lowy, I. et al. (1980) Cell 22:817-823.) Also, antimetabolite, antibiotic, or herbicide resistance can be used as the basis for selection. For example, *dhfr* confers resistance to methotrexate; *neo* confers resistance to the aminoglycosides neomycin and G-418; and *als* and *pat* confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively. (See, e.g., Wigler, M. et al. (1980) Proc. Natl. Acad. Sci. USA 77:3567-3570; Colbere-Garapin, F. et al. (1981) J. Mol. Biol. 150:1-14.) Additional selectable genes have been described, e.g., *trpB* and *hisD*, which alter cellular requirements for metabolites. (See, e.g., Hartman, S.C. and R.C. Mulligan (1988) Proc. Natl. Acad. Sci. USA 85:8047-8051.) Visible markers, e.g., anthocyanins, green fluorescent proteins (GFP; Clontech), ß glucuronidase and its substrate ß-glucuronide, or luciferase and its substrate luciferin may be used. These markers can be used not only to identify transformants, but also to quantify the amount of transient or stable protein expression attributable to a specific vector system. (See, e.g., Rhodes, C.A. (1995) Methods Mol. Biol. 55:121-131.)

Although the presence/absence of marker gene expression suggests that the gene of interest is also present, the presence and expression of the gene may need to be confirmed. For example, if the sequence encoding SECP is inserted within a marker gene sequence, transformed cells containing sequences encoding SECP can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a sequence encoding SECP under the control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the tandem gene as well.

In general, host cells that contain the nucleic acid sequence encoding SECP and that express SECP may be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations, PCR amplification, and protein bioassay or immunoassay techniques which include membrane, solution, or chip based technologies for the detection and/or quantification of nucleic acid or protein sequences.

Immunological methods for detecting and measuring the expression of SECP using either specific polyclonal or monoclonal antibodies are known in the art. Examples of such techniques include enzyme-linked immunosorbent assays (ELISAs), radioimmunoassays (RIAs), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on SECP is preferred, but a competitive binding assay may be employed. These and other assays are well known in the art. (See, e.g., Hampton, R. et al. (1990) Serological Methods, a Laboratory Manual, APS Press, St. Paul MN, Sect. IV; Coligan, J.E. et al. (1997) Current Protocols in Immunology, Greene Pub. Associates and Wiley-Interscience, New York NY; and Pound, J.D. (1998) Immunochemical Protocols, Humana Press, Totowa NJ.)

A wide variety of labels and conjugation techniques are known by those skilled in the art and

may be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to polynucleotides encoding SECP include oligolabeling, nick translation, end-labeling, or PCR amplification using a labeled nucleotide. Alternatively, the sequences encoding SECP, or any fragments thereof, may be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by addition of an appropriate RNA polymerase such as T7, T3, or SP6 and labeled nucleotides. These procedures may be conducted using a variety of commercially available kits, such as those provided by Amersham Pharmacia Biotech, Promega (Madison WI), and US Biochemical. Suitable reporter molecules or labels which may be used for ease of detection include radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents, as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

Host cells transformed with nucleotide sequences encoding SECP may be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The protein produced by a transformed cell may be secreted or retained intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides which encode SECP may be designed to contain signal sequences which direct secretion of SECP through a prokaryotic or eukaryotic cell membrane.

In addition, a host cell strain may be chosen for its ability to modulate expression of the inserted sequences or to process the expressed protein in the desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro" or "pro" form of the protein may also be used to specify protein targeting, folding, and/or activity. Different host cells which have specific cellular machinery and characteristic mechanisms for post-translational activities (e.g., CHO, HeLa, MDCK, HEK293, and WI38) are available from the American Type Culture Collection (ATCC, Manassas VA) and may be chosen to ensure the correct modification and processing of the foreign protein.

In another embodiment of the invention, natural, modified, or recombinant nucleic acid sequences encoding SECP may be ligated to a heterologous sequence resulting in translation of a fusion protein in any of the aforementioned host systems. For example, a chimeric SECP protein containing a heterologous moiety that can be recognized by a commercially available antibody may facilitate the screening of peptide libraries for inhibitors of SECP activity. Heterologous protein and peptide moieties may also facilitate purification of fusion proteins using commercially available affinity matrices. Such moieties include, but are not limited to, glutathione S-transferase (GST), maltose binding protein (MBP), thioredoxin (Trx), calmodulin binding peptide (CBP), 6-His, FLAG, *c-myc*, and hemagglutinin (HA). GST, MBP, Trx, CBP, and 6-His enable purification of their

cognate fusion proteins on immobilized glutathione, maltose, phenylarsine oxide, calmodulin, and metal-chelate resins, respectively. FLAG, *c-myc*, and hemagglutinin (HA) enable immunoaffinity purification of fusion proteins using commercially available monoclonal and polyclonal antibodies that specifically recognize these epitope tags. A fusion protein may also be engineered to contain a proteolytic cleavage site located between the SECP encoding sequence and the heterologous protein sequence, so that SECP may be cleaved away from the heterologous moiety following purification. Methods for fusion protein expression and purification are discussed in Ausubel (1995, supra, ch. 10). A variety of commercially available kits may also be used to facilitate expression and purification of fusion proteins.

In a further embodiment of the invention, synthesis of radiolabeled SECP may be achieved in vitro using the TNT rabbit reticulocyte lysate or wheat germ extract system (Promega). These systems couple transcription and translation of protein-coding sequences operably associated with the T7, T3, or SP6 promoters. Translation takes place in the presence of a radiolabeled amino acid precursor, for example, ^{35}S-methionine.

SECP of the present invention or fragments thereof may be used to screen for compounds that specifically bind to SECP. At least one and up to a plurality of test compounds may be screened for specific binding to SECP. Examples of test compounds include antibodies, oligonucleotides, proteins (e.g., receptors), or small molecules.

In one embodiment, the compound thus identified is closely related to the natural ligand of SECP, e.g., a ligand or fragment thereof, a natural substrate, a structural or functional mimetic, or a natural binding partner. (See, e.g., Coligan, J.E. et al. (1991) Current Protocols in Immunology 1(2): Chapter 5.) Similarly, the compound can be closely related to the natural receptor to which SECP binds, or to at least a fragment of the receptor, e.g., the ligand binding site. In either case, the compound can be rationally designed using known techniques. In one embodiment, screening for these compounds involves producing appropriate cells which express SECP, either as a secreted protein or on the cell membrane. Preferred cells include cells from mammals, yeast, Drosophila, or E. coli. Cells expressing SECP or cell membrane fractions which contain SECP are then contacted with a test compound and binding, stimulation, or inhibition of activity of either SECP or the compound is analyzed.

An assay may simply test binding of a test compound to the polypeptide, wherein binding is detected by a fluorophore, radioisotope, enzyme conjugate, or other detectable label. For example, the assay may comprise the steps of combining at least one test compound with SECP, either in solution or affixed to a solid support, and detecting the binding of SECP to the compound. Alternatively, the assay may detect or measure binding of a test compound in the presence of a labeled competitor. Additionally, the assay may be carried out using cell-free preparations, chemical

libraries, or natural product mixtures, and the test compound(s) may be free in solution or affixed to a solid support.

SECP of the present invention or fragments thereof may be used to screen for compounds that modulate the activity of SECP. Such compounds may include agonists, antagonists, or partial or inverse agonists. In one embodiment, an assay is performed under conditions permissive for SECP activity, wherein SECP is combined with at least one test compound, and the activity of SECP in the presence of a test compound is compared with the activity of SECP in the absence of the test compound. A change in the activity of SECP in the presence of the test compound is indicative of a compound that modulates the activity of SECP. Alternatively, a test compound is combined with an in vitro or cell-free system comprising SECP under conditions suitable for SECP activity, and the assay is performed. In either of these assays, a test compound which modulates the activity of SECP may do so indirectly and need not come in direct contact with the test compound. At least one and up to a plurality of test compounds may be screened.

In another embodiment, polynucleotides encoding SECP or their mammalian homologs may be "knocked out" in an animal model system using homologous recombination in embryonic stem (ES) cells. Such techniques are well known in the art and are useful for the generation of animal models of human disease. (See, e.g., U.S. Patent Number 5,175,383 and U.S. Patent Number 5,767,337.) For example, mouse ES cells, such as the mouse 129/SvJ cell line, are derived from the early mouse embryo and grown in culture. The ES cells are transformed with a vector containing the gene of interest disrupted by a marker gene, e.g., the neomycin phosphotransferase gene (neo; Capecchi, M.R. (1989) Science 244:1288-1292). The vector integrates into the corresponding region of the host genome by homologous recombination. Alternatively, homologous recombination takes place using the Cre-loxP system to knockout a gene of interest in a tissue- or developmental stage-specific manner (Marth, J.D. (1996) Clin. Invest. 97:1999-2002; Wagner, K.U. et al. (1997) Nucleic Acids Res. 25:4323-4330). Transformed ES cells are identified and microinjected into mouse cell blastocysts such as those from the C57BL/6 mouse strain. The blastocysts are surgically transferred to pseudopregnant dams, and the resulting chimeric progeny are genotyped and bred to produce heterozygous or homozygous strains. Transgenic animals thus generated may be tested with potential therapeutic or toxic agents.

Polynucleotides encoding SECP may also be manipulated in vitro in ES cells derived from human blastocysts. Human ES cells have the potential to differentiate into at least eight separate cell lineages including endoderm, mesoderm, and ectodermal cell types. These cell lineages differentiate into, for example, neural cells, hematopoietic lineages, and cardiomyocytes (Thomson, J.A. et al. (1998) Science 282:1145-1147).

Polynucleotides encoding SECP can also be used to create "knockin" humanized animals

(pigs) or transgenic animals (mice or rats) to model human disease. With knockin technology, a region of a polynucleotide encoding SECP is injected into animal ES cells, and the injected sequence integrates into the animal cell genome. Transformed cells are injected into blastulae, and the blastulae are implanted as described above. Transgenic progeny or inbred lines are studied and treated with potential pharmaceutical agents to obtain information on treatment of a human disease. Alternatively, a mammal inbred to overexpress SECP, e.g., by secreting SECP in its milk, may also serve as a convenient source of that protein (Janne, J. et al. (1998) Biotechnol. Annu. Rev. 4:55-74).

THERAPEUTICS

Chemical and structural similarity, e.g., in the context of sequences and motifs, exists between regions of SECP and secreted proteins. In addition, the expression of SECP is closely associated with reproductive, endocrine, immune system, gastrointestinal, fibroblastic, lung, brain and neurological tissue. Therefore, SECP appears to play a role in cell proliferative, autoimmune/inflammatory, cardiovascular, neurological, and developmental disorders. In the treatment of disorders associated with increased SECP expression or activity, it is desirable to decrease the expression or activity of SECP. In the treatment of disorders associated with decreased SECP expression or activity, it is desirable to increase the expression or activity of SECP.

Therefore, in one embodiment, SECP or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of SECP. Examples of such disorders include, but are not limited to, a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and cancers including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, a cancer of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; an autoimmune/inflammatory disorder such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's

syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma; a cardiovascular disorder such as congestive heart failure, ischemic heart disease, angina pectoris, myocardial infarction, hypertensive heart disease, degenerative valvular heart disease, calcific aortic valve stenosis, congenitally bicuspid aortic valve, mitral annular calcification, mitral valve prolapse, rheumatic fever and rheumatic heart disease, infective endocarditis, nonbacterial thrombotic endocarditis, endocarditis of systemic lupus erythematosus, carcinoid heart disease, cardiomyopathy, myocarditis, pericarditis, neoplastic heart disease, congenital heart disease, complications of cardiac transplantation, arteriovenous fistula, atherosclerosis, hypertension, vasculitis, Raynaud's disease, aneurysms, arterial dissections, varicose veins, thrombophlebitis and phlebothrombosis, vascular tumors, and complications of thrombolysis, balloon angioplasty, vascular replacement, and coronary artery bypass graft surgery; congenital lung anomalies, atelectasis, pulmonary congestion and edema, pulmonary embolism, pulmonary hemorrhage, pulmonary infarction, pulmonary hypertension, vascular sclerosis, obstructive pulmonary disease, restrictive pulmonary disease, chronic obstructive pulmonary disease, emphysema, chronic bronchitis, bronchial asthma, bronchiectasis, bacterial pneumonia, viral and mycoplasmal pneumonia, lung abscess, pulmonary tuberculosis, diffuse interstitial diseases, pneumoconioses, sarcoidosis, idiopathic pulmonary fibrosis, desquamative interstitial pneumonitis, hypersensitivity pneumonitis, pulmonary eosinophilia bronchiolitis obliterans-organizing pneumonia, diffuse pulmonary hemorrhage syndromes, Goodpasture's syndromes, idiopathic pulmonary hemosiderosis, pulmonary involvement in collagen-vascular disorders, pulmonary alveolar proteinosis, lung tumors, inflammatory and noninflammatory pleural effusions, pneumothorax, pleural tumors, drug-induced lung disease, radiation-induced lung disease, and complications of lung transplantation; a neurological disorder such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia, Parkinson's disease and other extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease, prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Straussler-Scheinker syndrome, fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorders of the central nervous system including Down syndrome, cerebral palsy, neuroskeletal disorders, autonomic

nervous system disorders, cranial nerve disorders, spinal cord diseases, muscular dystrophy and other neuromuscular disorders, peripheral nervous system disorders, dermatomyositis and polymyositis, inherited, metabolic, endocrine, and toxic myopathies, myasthenia gravis, periodic paralysis, mental disorders including mood, anxiety, and schizophrenic disorders, seasonal affective disorder (SAD), akathesia, amnesia, catatonia, diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic neuralgia, Tourette's disorder, progressive supranuclear palsy, corticobasal degeneration, and familial frontotemporal dementia and siezures; and a developmental disorder such as renal tubular acidosis, anemia, Cushing's syndrome, achondroplastic dwarfism, Duchenne and Becker muscular dystrophy, epilepsy, gonadal dysgenesis, WAGR syndrome (Wilms' tumor, aniridia, genitourinary abnormalities, and mental retardation), Smith-Magenis syndrome, myelodysplastic syndrome, hereditary mucoepithelial dysplasia, hereditary keratodermas, hereditary neuropathies such as Charcot-Marie-Tooth disease and neurofibromatosis, hypothyroidism, hydrocephalus, seizure disorders such as Syndenham's chorea and cerebral palsy, spina bifida, anencephaly, craniorachischisis, congenital glaucoma, cataract, and sensorineural hearing loss.

In another embodiment, a vector capable of expressing SECP or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of SECP including, but not limited to, those described above.

In a further embodiment, a composition comprising a substantially purified SECP in conjunction with a suitable pharmaceutical carrier may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of SECP including, but not limited to, those provided above.

In still another embodiment, an agonist which modulates the activity of SECP may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of SECP including, but not limited to, those listed above.

In a further embodiment, an antagonist of SECP may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of SECP. Examples of such disorders include, but are not limited to, those cell proliferative, autoimmune/inflammatory, cardiovascular, neurological, and developmental disorders described above. In one aspect, an antibody which specifically binds SECP may be used directly as an antagonist or indirectly as a targeting or delivery mechanism for bringing a pharmaceutical agent to cells or tissues which express SECP.

In an additional embodiment, a vector expressing the complement of the polynucleotide encoding SECP may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of SECP including, but not limited to, those described above.

In other embodiments, any of the proteins, antagonists, antibodies, agonists, complementary

sequences, or vectors of the invention may be administered in combination with other appropriate therapeutic agents. Selection of the appropriate agents for use in combination therapy may be made by one of ordinary skill in the art, according to conventional pharmaceutical principles. The combination of therapeutic agents may act synergistically to effect the treatment or prevention of the various disorders described above. Using this approach, one may be able to achieve therapeutic efficacy with lower dosages of each agent, thus reducing the potential for adverse side effects.

An antagonist of SECP may be produced using methods which are generally known in the art. In particular, purified SECP may be used to produce antibodies or to screen libraries of pharmaceutical agents to identify those which specifically bind SECP. Antibodies to SECP may also be generated using methods that are well known in the art. Such antibodies may include, but are not limited to, polyclonal, monoclonal, chimeric, and single chain antibodies, Fab fragments, and fragments produced by a Fab expression library. Neutralizing antibodies (i.e., those which inhibit dimer formation) are generally preferred for therapeutic use.

For the production of antibodies, various hosts including goats, rabbits, rats, mice, humans, and others may be immunized by injection with SECP or with any fragment or oligopeptide thereof which has immunogenic properties. Depending on the host species, various adjuvants may be used to increase immunological response. Such adjuvants include, but are not limited to, Freund's, mineral gels such as aluminum hydroxide, and surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, KLH, and dinitrophenol. Among adjuvants used in humans, BCG (bacilli Calmette-Guerin) and Corynebacterium parvum are especially preferable.

It is preferred that the oligopeptides, peptides, or fragments used to induce antibodies to SECP have an amino acid sequence consisting of at least about 5 amino acids, and generally will consist of at least about 10 amino acids. It is also preferable that these oligopeptides, peptides, or fragments are identical to a portion of the amino acid sequence of the natural protein. Short stretches of SECP amino acids may be fused with those of another protein, such as KLH, and antibodies to the chimeric molecule may be produced.

Monoclonal antibodies to SECP may be prepared using any technique which provides for the production of antibody molecules by continuous cell lines in culture. These include, but are not limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV-hybridoma technique. (See, e.g., Kohler, G. et al. (1975) Nature 256:495-497; Kozbor, D. et al. (1985) J. Immunol. Methods 81:31-42; Cote, R.J. et al. (1983) Proc. Natl. Acad. Sci. USA 80:2026-2030; and Cole, S.P. et al. (1984) Mol. Cell Biol. 62:109-120.)

In addition, techniques developed for the production of "chimeric antibodies," such as the splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity, can be used. (See, e.g., Morrison, S.L. et al. (1984) Proc.

Natl. Acad. Sci. USA 81:6851-6855; Neuberger, M.S. et al. (1984) Nature 312:604-608; and Takeda, S. et al. (1985) Nature 314:452-454.) Alternatively, techniques described for the production of single chain antibodies may be adapted, using methods known in the art, to produce SECP-specific single chain antibodies. Antibodies with related specificity, but of distinct idiotypic composition, may be generated by chain shuffling from random combinatorial immunoglobulin libraries. (See, e.g., Burton, D.R. (1991) Proc. Natl. Acad. Sci. USA 88:10134-10137.)

Antibodies may also be produced by inducing in vivo production in the lymphocyte population or by screening immunoglobulin libraries or panels of highly specific binding reagents as disclosed in the literature. (See, e.g., Orlandi, R. et al. (1989) Proc. Natl. Acad. Sci. USA 86:3833-3837; Winter, G. et al. (1991) Nature 349:293-299.)

Antibody fragments which contain specific binding sites for SECP may also be generated. For example, such fragments include, but are not limited to, $F(ab')_2$ fragments produced by pepsin digestion of the antibody molecule and Fab fragments generated by reducing the disulfide bridges of the F(ab')2 fragments. Alternatively, Fab expression libraries may be constructed to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity. (See, e.g., Huse, W.D. et al. (1989) Science 246:1275-1281.)

Various immunoassays may be used for screening to identify antibodies having the desired specificity. Numerous protocols for competitive binding or immunoradiometric assays using either polyclonal or monoclonal antibodies with established specificities are well known in the art. Such immunoassays typically involve the measurement of complex formation between SECP and its specific antibody. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering SECP epitopes is generally used, but a competitive binding assay may also be employed (Pound, supra).

Various methods such as Scatchard analysis in conjunction with radioimmunoassay techniques may be used to assess the affinity of antibodies for SECP. Affinity is expressed as an association constant, K_a, which is defined as the molar concentration of SECP-antibody complex divided by the molar concentrations of free antigen and free antibody under equilibrium conditions. The K_a determined for a preparation of polyclonal antibodies, which are heterogeneous in their affinities for multiple SECP epitopes, represents the average affinity, or avidity, of the antibodies for SECP. The K_a determined for a preparation of monoclonal antibodies, which are monospecific for a particular SECP epitope, represents a true measure of affinity. High-affinity antibody preparations with K_a ranging from about 10^9 to 10^{12} L/mole are preferred for use in immunoassays in which the SECP-antibody complex must withstand rigorous manipulations. Low-affinity antibody preparations with K_a ranging from about 10^6 to 10^7 L/mole are preferred for use in immunopurification and similar procedures which ultimately require dissociation of SECP, preferably in active form, from the

antibody (Catty, D. (1988) Antibodies, Volume I: A Practical Approach, IRL Press, Washington DC; Liddell, J.E. and A. Cryer (1991) A Practical Guide to Monoclonal Antibodies, John Wiley & Sons, New York NY).

The titer and avidity of polyclonal antibody preparations may be further evaluated to determine the quality and suitability of such preparations for certain downstream applications. For example, a polyclonal antibody preparation containing at least 1-2 mg specific antibody/ml, preferably 5-10 mg specific antibody/ml, is generally employed in procedures requiring precipitation of SECP-antibody complexes. Procedures for evaluating antibody specificity, titer, and avidity, and guidelines for antibody quality and usage in various applications, are generally available. (See, e.g., Catty, supra, and Coligan et al. supra.)

In another embodiment of the invention, the polynucleotides encoding SECP, or any fragment or complement thereof, may be used for therapeutic purposes. In one aspect, modifications of gene expression can be achieved by designing complementary sequences or antisense molecules (DNA, RNA, PNA, or modified oligonucleotides) to the coding or regulatory regions of the gene encoding SECP. Such technology is well known in the art, and antisense oligonucleotides or larger fragments can be designed from various locations along the coding or control regions of sequences encoding SECP. (See, e.g., Agrawal, S., ed. (1996) Antisense Therapeutics, Humana Press Inc., Totawa NJ.)

In therapeutic use, any gene delivery system suitable for introduction of the antisense sequences into appropriate target cells can be used. Antisense sequences can be delivered intracellularly in the form of an expression plasmid which, upon transcription, produces a sequence complementary to at least a portion of the cellular sequence encoding the target protein. (See, e.g., Slater, J.E. et al. (1998) J. Allergy Cli. Immunol. 102(3):469-475; and Scanlon, K.J. et al. (1995) 9(13):1288-1296.) Antisense sequences can also be introduced intracellularly through the use of viral vectors, such as retrovirus and adeno-associated virus vectors. (See, e.g., Miller, A.D. (1990) Blood 76:271; Ausubel, supra; Uckert, W. and W. Walther (1994) Pharmacol. Ther. 63(3):323-347.) Other gene delivery mechanisms include liposome-derived systems, artificial viral envelopes, and other systems known in the art. (See, e.g., Rossi, J.J. (1995) Br. Med. Bull. 51(1):217-225; Boado, R.J. et al. (1998) J. Pharm. Sci. 87(11):1308-1315; and Morris, M.C. et al. (1997) Nucleic Acids Res. 25(14):2730-2736.)

In another embodiment of the invention, polynucleotides encoding SECP may be used for somatic or germline gene therapy. Gene therapy may be performed to (i) correct a genetic deficiency (e.g., in the cases of severe combined immunodeficiency (SCID)-X1 disease characterized by X-linked inheritance (Cavazzana-Calvo, M. et al. (2000) Science 288:669-672), severe combined immunodeficiency syndrome associated with an inherited adenosine deaminase (ADA) deficiency (Blaese, R.M. et al. (1995) Science 270:475-480; Bordignon, C. et al. (1995) Science 270:470-475),

cystic fibrosis (Zabner, J. et al. (1993) Cell 75:207-216; Crystal, R.G. et al. (1995) Hum. Gene Therapy 6:643-666; Crystal, R.G. et al. (1995) Hum. Gene Therapy 6:667-703), thalassamias, familial hypercholesterolemia, and hemophilia resulting from Factor VIII or Factor IX deficiencies (Crystal, R.G. (1995) Science 270:404-410; Verma, I.M. and N. Somia (1997) Nature 389:239-242)), (ii) express a conditionally lethal gene product (e.g., in the case of cancers which result from unregulated cell proliferation), or (iii) express a protein which affords protection against intracellular parasites (e.g., against human retroviruses, such as human immunodeficiency virus (HIV) (Baltimore, D. (1988) Nature 335:395-396; Poeschla, E. et al. (1996) Proc. Natl. Acad. Sci. USA. 93:11395-11399), hepatitis B or C virus (HBV, HCV); fungal parasites, such as Candida albicans and Paracoccidioides brasiliensis; and protozoan parasites such as Plasmodium falciparum and Trypanosoma cruzi). In the case where a genetic deficiency in SECP expression or regulation causes disease, the expression of SECP from an appropriate population of transduced cells may alleviate the clinical manifestations caused by the genetic deficiency.

In a further embodiment of the invention, diseases or disorders caused by deficiencies in SECP are treated by constructing mammalian expression vectors encoding SECP and introducing these vectors by mechanical means into SECP-deficient cells. Mechanical transfer technologies for use with cells in vivo or ex vitro include (i) direct DNA microinjection into individual cells, (ii) ballistic gold particle delivery, (iii) liposome-mediated transfection, (iv) receptor-mediated gene transfer, and (v) the use of DNA transposons (Morgan, R.A. and W.F. Anderson (1993) Annu. Rev. Biochem. 62:191-217; Ivics, Z. (1997) Cell 91:501-510; Boulay, J-L. and H. Récipon (1998) Curr. Opin. Biotechnol. 9:445-450).

Expression vectors that may be effective for the expression of SECP include, but are not limited to, the PCDNA 3.1, EPITAG, PRCCMV2, PREP, PVAX vectors (Invitrogen, Carlsbad CA), PCMV-SCRIPT, PCMV-TAG, PEGSH/PERV (Stratagene, La Jolla CA), and PTET-OFF, PTET-ON, PTRE2, PTRE2-LUC, PTK-HYG (Clontech, Palo Alto CA). SECP may be expressed using (i) a constitutively active promoter, (e.g., from cytomegalovirus (CMV), Rous sarcoma virus (RSV), SV40 virus, thymidine kinase (TK), or β-actin genes), (ii) an inducible promoter (e.g., the tetracycline-regulated promoter (Gossen, M. and H. Bujard (1992) Proc. Natl. Acad. Sci. USA 89:5547-5551; Gossen, M. et al. (1995) Science 268:1766-1769; Rossi, F.M.V. and H.M. Blau (1998) Curr. Opin. Biotechnol. 9:451-456), commercially available in the T-REX plasmid (Invitrogen)); the ecdysone-inducible promoter (available in the plasmids PVGRXR and PIND; Invitrogen); the FK506/rapamycin inducible promoter; or the RU486/mifepristone inducible promoter (Rossi, F.M.V. and Blau, H.M. supra)), or (iii) a tissue-specific promoter or the native promoter of the endogenous gene encoding SECP from a normal individual.

Commercially available liposome transformation kits (e.g., the PERFECT LIPID

TRANSFECTION KIT, available from Invitrogen) allow one with ordinary skill in the art to deliver polynucleotides to target cells in culture and require minimal effort to optimize experimental parameters. In the alternative, transformation is performed using the calcium phosphate method (Graham, F.L. and A.J. Eb (1973) Virology 52:456-467), or by electroporation (Neumann, E. et al. (1982) EMBO J. 1:841-845). The introduction of DNA to primary cells requires modification of these standardized mammalian transfection protocols.

In another embodiment of the invention, diseases or disorders caused by genetic defects with respect to SECP expression are treated by constructing a retrovirus vector consisting of (i) the polynucleotide encoding SECP under the control of an independent promoter or the retrovirus long terminal repeat (LTR) promoter, (ii) appropriate RNA packaging signals, and (iii) a Rev-responsive element (RRE) along with additional retrovirus *cis*-acting RNA sequences and coding sequences required for efficient vector propagation. Retrovirus vectors (e.g., PFB and PFBNEO) are commercially available (Stratagene) and are based on published data (Riviere, I. et al. (1995) Proc. Natl. Acad. Sci. USA 92:6733-6737), incorporated by reference herein. The vector is propagated in an appropriate vector producing cell line (VPCL) that expresses an envelope gene with a tropism for receptors on the target cells or a promiscuous envelope protein such as VSVg (Armentano, D. et al. (1987) J. Virol. 61:1647-1650; Bender, M.A. et al. (1987) J. Virol. 61:1639-1646; Adam, M.A. and A.D. Miller (1988) J. Virol. 62:3802-3806; Dull, T. et al. (1998) J. Virol. 72:8463-8471; Zufferey, R. et al. (1998) J. Virol. 72:9873-9880). U.S. Patent Number 5,910,434 to Rigg ("Method for obtaining retrovirus packaging cell lines producing high transducing efficiency retroviral supernatant") discloses a method for obtaining retrovirus packaging cell lines and is hereby incorporated by reference. Propagation of retrovirus vectors, transduction of a population of cells (e.g., $CD4^+$ T-cells), and the return of transduced cells to a patient are procedures well known to persons skilled in the art of gene therapy and have been well documented (Ranga, U. et al. (1997) J. Virol. 71:7020-7029; Bauer, G. et al. (1997) Blood 89:2259-2267; Bonyhadi, M.L. (1997) J. Virol. 71:4707-4716; Ranga, U. et al. (1998) Proc. Natl. Acad. Sci. USA 95:1201-1206; Su, L. (1997) Blood 89:2283-2290).

In the alternative, an adenovirus-based gene therapy delivery system is used to deliver polynucleotides encoding SECP to cells which have one or more genetic abnormalities with respect to the expression of SECP. The construction and packaging of adenovirus-based vectors are well known to those with ordinary skill in the art. Replication defective adenovirus vectors have proven to be versatile for importing genes encoding immunoregulatory proteins into intact islets in the pancreas (Csete, M.E. et al. (1995) Transplantation 27:263-268). Potentially useful adenoviral vectors are described in U.S. Patent Number 5,707,618 to Armentano ("Adenovirus vectors for gene therapy"), hereby incorporated by reference. For adenoviral vectors, see also Antinozzi, P.A. et al. (1999)

Annu. Rev. Nutr. 19:511-544 and Verma, I.M. and N. Somia (1997) Nature 18:389:239-242, both incorporated by reference herein.

In another alternative, a herpes-based, gene therapy delivery system is used to deliver polynucleotides encoding SECP to target cells which have one or more genetic abnormalities with respect to the expression of SECP. The use of herpes simplex virus (HSV)-based vectors may be especially valuable for introducing SECP to cells of the central nervous system, for which HSV has a tropism. The construction and packaging of herpes-based vectors are well known to those with ordinary skill in the art. A replication-competent herpes simplex virus (HSV) type 1-based vector has been used to deliver a reporter gene to the eyes of primates (Liu, X. et al. (1999) Exp. Eye Res. 169:385-395). The construction of a HSV-1 virus vector has also been disclosed in detail in U.S. Patent Number 5,804,413 to DeLuca ("Herpes simplex virus strains for gene transfer"), which is hereby incorporated by reference. U.S. Patent Number 5,804,413 teaches the use of recombinant HSV d92 which consists of a genome containing at least one exogenous gene to be transferred to a cell under the control of the appropriate promoter for purposes including human gene therapy. Also taught by this patent are the construction and use of recombinant HSV strains deleted for ICP4, ICP27 and ICP22. For HSV vectors, see also Goins, W.F. et al. (1999) J. Virol. 73:519-532 and Xu, H. et al. (1994) Dev. Biol. 163:152-161, hereby incorporated by reference. The manipulation of cloned herpesvirus sequences, the generation of recombinant virus following the transfection of multiple plasmids containing different segments of the large herpesvirus genomes, the growth and propagation of herpesvirus, and the infection of cells with herpesvirus are techniques well known to those of ordinary skill in the art.

In another alternative, an alphavirus (positive, single-stranded RNA virus) vector is used to deliver polynucleotides encoding SECP to target cells. The biology of the prototypic alphavirus, Semliki Forest Virus (SFV), has been studied extensively and gene transfer vectors have been based on the SFV genome (Garoff, H. and K.-J. Li (1998) Curr. Opin. Biotechnol. 9:464-469). During alphavirus RNA replication, a subgenomic RNA is generated that normally encodes the viral capsid proteins. This subgenomic RNA replicates to higher levels than the full length genomic RNA, resulting in the overproduction of capsid proteins relative to the viral proteins with enzymatic activity (e.g., protease and polymerase). Similarly, inserting the coding sequence for SECP into the alphavirus genome in place of the capsid-coding region results in the production of a large number of SECP-coding RNAs and the synthesis of high levels of SECP in vector transduced cells. While alphavirus infection is typically associated with cell lysis within a few days, the ability to establish a persistent infection in hamster normal kidney cells (BHK-21) with a variant of Sindbis virus (SIN) indicates that the lytic replication of alphaviruses can be altered to suit the needs of the gene therapy application (Dryga, S.A. et al. (1997) Virology 228:74-83). The wide host range of alphaviruses will

allow the introduction of SECP into a variety of cell types. The specific transduction of a subset of cells in a population may require the sorting of cells prior to transduction. The methods of manipulating infectious cDNA clones of alphaviruses, performing alphavirus cDNA and RNA transfections, and performing alphavirus infections, are well known to those with ordinary skill in the art.

Oligonucleotides derived from the transcription initiation site, e.g., between about positions -10 and +10 from the start site, may also be employed to inhibit gene expression. Similarly, inhibition can be achieved using triple helix base-pairing methodology. Triple helix pairing is useful because it causes inhibition of the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or regulatory molecules. Recent therapeutic advances using triplex DNA have been described in the literature. (See, e.g., Gee, J.E. et al. (1994) in Huber, B.E. and B.I. Carr, Molecular and Immunologic Approaches, Futura Publishing, Mt. Kisco NY, pp. 163-177.) A complementary sequence or antisense molecule may also be designed to block translation of mRNA by preventing the transcript from binding to ribosomes.

Ribozymes, enzymatic RNA molecules, may also be used to catalyze the specific cleavage of RNA. The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme molecule to complementary target RNA, followed by endonucleolytic cleavage. For example, engineered hammerhead motif ribozyme molecules may specifically and efficiently catalyze endonucleolytic cleavage of sequences encoding SECP.

Specific ribozyme cleavage sites within any potential RNA target are initially identified by scanning the target molecule for ribozyme cleavage sites, including the following sequences: GUA, GUU, and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides, corresponding to the region of the target gene containing the cleavage site, may be evaluated for secondary structural features which may render the oligonucleotide inoperable. The suitability of candidate targets may also be evaluated by testing accessibility to hybridization with complementary oligonucleotides using ribonuclease protection assays.

Complementary ribonucleic acid molecules and ribozymes of the invention may be prepared by any method known in the art for the synthesis of nucleic acid molecules. These include techniques for chemically synthesizing oligonucleotides such as solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by in vitro and in vivo transcription of DNA sequences encoding SECP. Such DNA sequences may be incorporated into a wide variety of vectors with suitable RNA polymerase promoters such as T7 or SP6. Alternatively, these cDNA constructs that synthesize complementary RNA, constitutively or inducibly, can be introduced into cell lines, cells, or tissues.

RNA molecules may be modified to increase intracellular stability and half-life. Possible

modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends of the molecule, or the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages within the backbone of the molecule. This concept is inherent in the production of PNAs and can be extended in all of these molecules by the inclusion of nontraditional bases such as inosine, queosine, and wybutosine, as well as acetyl-, methyl-, thio-, and similarly modified forms of adenine, cytidine, guanine, thymine, and uridine which are not as easily recognized by endogenous endonucleases.

An additional embodiment of the invention encompasses a method for screening for a compound which is effective in altering expression of a polynucleotide encoding SECP. Compounds which may be effective in altering expression of a specific polynucleotide may include, but are not limited to, oligonucleotides, antisense oligonucleotides, triple helix-forming oligonucleotides, transcription factors and other polypeptide transcriptional regulators, and non-macromolecular chemical entities which are capable of interacting with specific polynucleotide sequences. Effective compounds may alter polynucleotide expression by acting as either inhibitors or promoters of polynucleotide expression. Thus, in the treatment of disorders associated with increased SECP expression or activity, a compound which specifically inhibits expression of the polynucleotide encoding SECP may be therapeutically useful, and in the treatment of disorders associated with decreased SECP expression or activity, a compound which specifically promotes expression of the polynucleotide encoding SECP may be therapeutically useful.

At least one, and up to a plurality, of test compounds may be screened for effectiveness in altering expression of a specific polynucleotide. A test compound may be obtained by any method commonly known in the art, including chemical modification of a compound known to be effective in altering polynucleotide expression; selection from an existing, commercially-available or proprietary library of naturally-occurring or non-natural chemical compounds; rational design of a compound based on chemical and/or structural properties of the target polynucleotide; and selection from a library of chemical compounds created combinatorially or randomly. A sample comprising a polynucleotide encoding SECP is exposed to at least one test compound thus obtained. The sample may comprise, for example, an intact or permeabilized cell, or an in vitro cell-free or reconstituted biochemical system. Alterations in the expression of a polynucleotide encoding SECP are assayed by any method commonly known in the art. Typically, the expression of a specific nucleotide is detected by hybridization with a probe having a nucleotide sequence complementary to the sequence of the polynucleotide encoding SECP. The amount of hybridization may be quantified, thus forming the basis for a comparison of the expression of the polynucleotide both with and without exposure to one or more test compounds. Detection of a change in the expression of a polynucleotide exposed to a test compound indicates that the test compound is effective in altering the expression of the

polynucleotide. A screen for a compound effective in altering expression of a specific polynucleotide can be carried out, for example, using a Schizosaccharomyces pombe gene expression system (Atkins, D. et al. (1999) U.S. Patent No. 5,932,435; Arndt, G.M. et al. (2000) Nucleic Acids Res. 28:E15) or a human cell line such as HeLa cell (Clarke, M.L. et al. (2000) Biochem. Biophys. Res. Commun. 268:8-13). A particular embodiment of the present invention involves screening a combinatorial library of oligonucleotides (such as deoxyribonucleotides, ribonucleotides, peptide nucleic acids, and modified oligonucleotides) for antisense activity against a specific polynucleotide sequence (Bruice, T.W. et al. (1997) U.S. Patent No. 5,686,242; Bruice, T.W. et al. (2000) U.S. Patent No. 6,022,691).

Many methods for introducing vectors into cells or tissues are available and equally suitable for use in vivo, in vitro, and ex vivo. For ex vivo therapy, vectors may be introduced into stem cells taken from the patient and clonally propagated for autologous transplant back into that same patient. Delivery by transfection, by liposome injections, or by polycationic amino polymers may be achieved using methods which are well known in the art. (See, e.g., Goldman, C.K. et al. (1997) Nat. Biotechnol. 15:462-466.)

Any of the therapeutic methods described above may be applied to any subject in need of such therapy, including, for example, mammals such as humans, dogs, cats, cows, horses, rabbits, and monkeys.

An additional embodiment of the invention relates to the administration of a composition which generally comprises an active ingredient formulated with a pharmaceutically acceptable excipient. Excipients may include, for example, sugars, starches, celluloses, gums, and proteins. Various formulations are commonly known and are thoroughly discussed in the latest edition of Remington's Pharmaceutical Sciences (Maack Publishing, Easton PA). Such compositions may consist of SECP, antibodies to SECP, and mimetics, agonists, antagonists, or inhibitors of SECP.

The compositions utilized in this invention may be administered by any number of routes including, but not limited to, oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, pulmonary, transdermal, subcutaneous, intraperitoneal, intranasal, enteral, topical, sublingual, or rectal means.

Compositions for pulmonary administration may be prepared in liquid or dry powder form. These compositions are generally aerosolized immediately prior to inhalation by the patient. In the case of small molecules (e.g. traditional low molecular weight organic drugs), aerosol delivery of fast-acting formulations is well-known in the art. In the case of macromolecules (e.g. larger peptides and proteins), recent developments in the field of pulmonary delivery via the alveolar region of the lung have enabled the practical delivery of drugs such as insulin to blood circulation (see, e.g., Patton, J.S. et al., U.S. Patent No. 5,997,848). Pulmonary delivery has the advantage of administration

without needle injection, and obviates the need for potentially toxic penetration enhancers.

Compositions suitable for use in the invention include compositions wherein the active ingredients are contained in an effective amount to achieve the intended purpose. The determination of an effective dose is well within the capability of those skilled in the art.

Specialized forms of compositions may be prepared for direct intracellular delivery of macromolecules comprising SECP or fragments thereof. For example, liposome preparations containing a cell-impermeable macromolecule may promote cell fusion and intracellular delivery of the macromolecule. Alternatively, SECP or a fragment thereof may be joined to a short cationic N-terminal portion from the HIV Tat-1 protein. Fusion proteins thus generated have been found to transduce into the cells of all tissues, including the brain, in a mouse model system (Schwarze, S.R. et al. (1999) Science 285:1569-1572).

For any compound, the therapeutically effective dose can be estimated initially either in cell culture assays, e.g., of neoplastic cells, or in animal models such as mice, rats, rabbits, dogs, monkeys, or pigs. An animal model may also be used to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans.

A therapeutically effective dose refers to that amount of active ingredient, for example SECP or fragments thereof, antibodies of SECP, and agonists, antagonists or inhibitors of SECP, which ameliorates the symptoms or condition. Therapeutic efficacy and toxicity may be determined by standard pharmaceutical procedures in cell cultures or with experimental animals, such as by calculating the ED_{50} (the dose therapeutically effective in 50% of the population) or LD_{50} (the dose lethal to 50% of the population) statistics. The dose ratio of toxic to therapeutic effects is the therapeutic index, which can be expressed as the LD_{50}/ED_{50} ratio. Compositions which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies are used to formulate a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that includes the ED_{50} with little or no toxicity. The dosage varies within this range depending upon the dosage form employed, the sensitivity of the patient, and the route of administration.

The exact dosage will be determined by the practitioner, in light of factors related to the subject requiring treatment. Dosage and administration are adjusted to provide sufficient levels of the active moiety or to maintain the desired effect. Factors which may be taken into account include the severity of the disease state, the general health of the subject, the age, weight, and gender of the subject, time and frequency of administration, drug combination(s), reaction sensitivities, and response to therapy. Long-acting compositions may be administered every 3 to 4 days, every week, or biweekly depending on the half-life and clearance rate of the particular formulation.

Normal dosage amounts may vary from about 0.1 μg to 100,000 μg, up to a total dose of about 1 gram, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature and generally available to practitioners in the art. Those skilled in the art will employ different formulations for nucleotides than for proteins or their inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc.

DIAGNOSTICS

In another embodiment, antibodies which specifically bind SECP may be used for the diagnosis of disorders characterized by expression of SECP, or in assays to monitor patients being treated with SECP or agonists, antagonists, or inhibitors of SECP. Antibodies useful for diagnostic purposes may be prepared in the same manner as described above for therapeutics. Diagnostic assays for SECP include methods which utilize the antibody and a label to detect SECP in human body fluids or in extracts of cells or tissues. The antibodies may be used with or without modification, and may be labeled by covalent or non-covalent attachment of a reporter molecule. A wide variety of reporter molecules, several of which are described above, are known in the art and may be used.

A variety of protocols for measuring SECP, including ELISAs, RIAs, and FACS, are known in the art and provide a basis for diagnosing altered or abnormal levels of SECP expression. Normal or standard values for SECP expression are established by combining body fluids or cell extracts taken from normal mammalian subjects, for example, human subjects, with antibodies to SECP under conditions suitable for complex formation. The amount of standard complex formation may be quantitated by various methods, such as photometric means. Quantities of SECP expressed in subject, control, and disease samples from biopsied tissues are compared with the standard values. Deviation between standard and subject values establishes the parameters for diagnosing disease.

In another embodiment of the invention, the polynucleotides encoding SECP may be used for diagnostic purposes. The polynucleotides which may be used include oligonucleotide sequences, complementary RNA and DNA molecules, and PNAs. The polynucleotides may be used to detect and quantify gene expression in biopsied tissues in which expression of SECP may be correlated with disease. The diagnostic assay may be used to determine absence, presence, and excess expression of SECP, and to monitor regulation of SECP levels during therapeutic intervention.

In one aspect, hybridization with PCR probes which are capable of detecting polynucleotide sequences, including genomic sequences, encoding SECP or closely related molecules may be used to identify nucleic acid sequences which encode SECP. The specificity of the probe, whether it is made from a highly specific region, e.g., the 5' regulatory region, or from a less specific region, e.g., a conserved motif, and the stringency of the hybridization or amplification will determine whether the probe identifies only naturally occurring sequences encoding SECP, allelic variants, or related

sequences.

Probes may also be used for the detection of related sequences, and may have at least 50% sequence identity to any of the SECP encoding sequences. The hybridization probes of the subject invention may be DNA or RNA and may be derived from the sequence of SEQ ID NO:45-88 or from genomic sequences including promoters, enhancers, and introns of the SECP gene.

Means for producing specific hybridization probes for DNAs encoding SECP include the cloning of polynucleotide sequences encoding SECP or SECP derivatives into vectors for the production of mRNA probes. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by means of the addition of the appropriate RNA polymerases and the appropriate labeled nucleotides. Hybridization probes may be labeled by a variety of reporter groups, for example, by radionuclides such as ^{32}P or ^{35}S, or by enzymatic labels, such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems, and the like.

Polynucleotide sequences encoding SECP may be used for the diagnosis of disorders associated with expression of SECP. Examples of such disorders include, but are not limited to, a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and cancers including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, a cancer of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; an autoimmune/inflammatory disorder such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma; a cardiovascular disorder such as congestive heart failure, ischemic heart disease, angina pectoris,

myocardial infarction, hypertensive heart disease, degenerative valvular heart disease, calcific aortic valve stenosis, congenitally bicuspid aortic valve, mitral annular calcification, mitral valve prolapse, rheumatic fever and rheumatic heart disease, infective endocarditis, nonbacterial thrombotic endocarditis, endocarditis of systemic lupus erythematosus, carcinoid heart disease, cardiomyopathy, myocarditis, pericarditis, neoplastic heart disease, congenital heart disease, complications of cardiac transplantation, arteriovenous fistula, atherosclerosis, hypertension, vasculitis, Raynaud's disease, aneurysms, arterial dissections, varicose veins, thrombophlebitis and phlebothrombosis, vascular tumors, and complications of thrombolysis, balloon angioplasty, vascular replacement, and coronary artery bypass graft surgery; congenital lung anomalies, atelectasis, pulmonary congestion and edema, pulmonary embolism, pulmonary hemorrhage, pulmonary infarction, pulmonary hypertension, vascular sclerosis, obstructive pulmonary disease, restrictive pulmonary disease, chronic obstructive pulmonary disease, emphysema, chronic bronchitis, bronchial asthma, bronchiectasis, bacterial pneumonia, viral and mycoplasmal pneumonia, lung abscess, pulmonary tuberculosis, diffuse interstitial diseases, pneumoconioses, sarcoidosis, idiopathic pulmonary fibrosis, desquamative interstitial pneumonitis, hypersensitivity pneumonitis, pulmonary eosinophilia bronchiolitis obliterans-organizing pneumonia, diffuse pulmonary hemorrhage syndromes, Goodpasture's syndromes, idiopathic pulmonary hemosiderosis, pulmonary involvement in collagen-vascular disorders, pulmonary alveolar proteinosis, lung tumors, inflammatory and noninflammatory pleural effusions, pneumothorax, pleural tumors, drug-induced lung disease, radiation-induced lung disease, and complications of lung transplantation; a neurological disorder such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia, Parkinson's disease and other extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease, prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Straussler-Scheinker syndrome, fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorders of the central nervous system including Down syndrome, cerebral palsy, neuroskeletal disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord diseases, muscular dystrophy and other neuromuscular disorders, peripheral nervous system disorders, dermatomyositis and polymyositis, inherited, metabolic, endocrine, and toxic myopathies, myasthenia gravis, periodic paralysis, mental disorders including mood, anxiety, and schizophrenic disorders, seasonal affective disorder (SAD), akathesia, amnesia, catatonia,

diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic neuralgia, Tourette's disorder, progressive supranuclear palsy, corticobasal degeneration, and familial frontotemporal dementia and siezures; and a developmental disorder such as renal tubular acidosis, anemia, Cushing's syndrome, achondroplastic dwarfism, Duchenne and Becker muscular dystrophy, epilepsy, gonadal dysgenesis, WAGR syndrome (Wilms' tumor, aniridia, genitourinary abnormalities, and mental retardation), Smith-Magenis syndrome, myelodysplastic syndrome, hereditary mucoepithelial dysplasia, hereditary keratodermas, hereditary neuropathies such as Charcot-Marie-Tooth disease and neurofibromatosis, hypothyroidism, hydrocephalus, seizure disorders such as Syndenham's chorea and cerebral palsy, spina bifida, anencephaly, craniorachischisis, congenital glaucoma, cataract, and sensorineural hearing loss. The polynucleotide sequences encoding SECP may be used in Southern or northern analysis, dot blot, or other membrane-based technologies; in PCR technologies; in dipstick, pin, and multiformat ELISA-like assays; and in microarrays utilizing fluids or tissues from patients to detect altered SECP expression. Such qualitative or quantitative methods are well known in the art.

In a particular aspect, the nucleotide sequences encoding SECP may be useful in assays that detect the presence of associated disorders, particularly those mentioned above. The nucleotide sequences encoding SECP may be labeled by standard methods and added to a fluid or tissue sample from a patient under conditions suitable for the formation of hybridization complexes. After a suitable incubation period, the sample is washed and the signal is quantified and compared with a standard value. If the amount of signal in the patient sample is significantly altered in comparison to a control sample then the presence of altered levels of nucleotide sequences encoding SECP in the sample indicates the presence of the associated disorder. Such assays may also be used to evaluate the efficacy of a particular therapeutic treatment regimen in animal studies, in clinical trials, or to monitor the treatment of an individual patient.

In order to provide a basis for the diagnosis of a disorder associated with expression of SECP, a normal or standard profile for expression is established. This may be accomplished by combining body fluids or cell extracts taken from normal subjects, either animal or human, with a sequence, or a fragment thereof, encoding SECP, under conditions suitable for hybridization or amplification. Standard hybridization may be quantified by comparing the values obtained from normal subjects with values from an experiment in which a known amount of a substantially purified polynucleotide is used. Standard values obtained in this manner may be compared with values obtained from samples from patients who are symptomatic for a disorder. Deviation from standard values is used to establish the presence of a disorder.

Once the presence of a disorder is established and a treatment protocol is initiated, hybridization assays may be repeated on a regular basis to determine if the level of expression in the

patient begins to approximate that which is observed in the normal subject. The results obtained from successive assays may be used to show the efficacy of treatment over a period ranging from several days to months.

With respect to cancer, the presence of an abnormal amount of transcript (either under- or overexpressed) in biopsied tissue from an individual may indicate a predisposition for the development of the disease, or may provide a means for detecting the disease prior to the appearance of actual clinical symptoms. A more definitive diagnosis of this type may allow health professionals to employ preventative measures or aggressive treatment earlier thereby preventing the development or further progression of the cancer.

Additional diagnostic uses for oligonucleotides designed from the sequences encoding SECP may involve the use of PCR. These oligomers may be chemically synthesized, generated enzymatically, or produced in vitro. Oligomers will preferably contain a fragment of a polynucleotide encoding SECP, or a fragment of a polynucleotide complementary to the polynucleotide encoding SECP, and will be employed under optimized conditions for identification of a specific gene or condition. Oligomers may also be employed under less stringent conditions for detection or quantification of closely related DNA or RNA sequences.

In a particular aspect, oligonucleotide primers derived from the polynucleotide sequences encoding SECP may be used to detect single nucleotide polymorphisms (SNPs). SNPs are substitutions, insertions and deletions that are a frequent cause of inherited or acquired genetic disease in humans. Methods of SNP detection include, but are not limited to, single-stranded conformation polymorphism (SSCP) and fluorescent SSCP (fSSCP) methods. In SSCP, oligonucleotide primers derived from the polynucleotide sequences encoding SECP are used to amplify DNA using the polymerase chain reaction (PCR). The DNA may be derived, for example, from diseased or normal tissue, biopsy samples, bodily fluids, and the like. SNPs in the DNA cause differences in the secondary and tertiary structures of PCR products in single-stranded form, and these differences are detectable using gel electrophoresis in non-denaturing gels. In fSCCP, the oligonucleotide primers are fluorescently labeled, which allows detection of the amplimers in high-throughput equipment such as DNA sequencing machines. Additionally, sequence database analysis methods, termed in silico SNP (isSNP), are capable of identifying polymorphisms by comparing the sequence of individual overlapping DNA fragments which assemble into a common consensus sequence. These computer-based methods filter out sequence variations due to laboratory preparation of DNA and sequencing errors using statistical models and automated analyses of DNA sequence chromatograms. In the alternative, SNPs may be detected and characterized by mass spectrometry using, for example, the high throughput MASSARRAY system (Sequenom, Inc., San Diego CA).

Methods which may also be used to quantify the expression of SECP include radiolabeling or

biotinylating nucleotides, coamplification of a control nucleic acid, and interpolating results from standard curves. (See, e.g., Melby, P.C. et al. (1993) J. Immunol. Methods 159:235-244; Duplaa, C. et al. (1993) Anal. Biochem. 212:229-236.) The speed of quantitation of multiple samples may be accelerated by running the assay in a high-throughput format where the oligomer or polynucleotide of interest is presented in various dilutions and a spectrophotometric or colorimetric response gives rapid quantitation.

In further embodiments, oligonucleotides or longer fragments derived from any of the polynucleotide sequences described herein may be used as elements on a microarray. The microarray can be used in transcript imaging techniques which monitor the relative expression levels of large numbers of genes simultaneously as described below. The microarray may also be used to identify genetic variants, mutations, and polymorphisms. This information may be used to determine gene function, to understand the genetic basis of a disorder, to diagnose a disorder, to monitor progression/regression of disease as a function of gene expression, and to develop and monitor the activities of therapeutic agents in the treatment of disease. In particular, this information may be used to develop a pharmacogenomic profile of a patient in order to select the most appropriate and effective treatment regimen for that patient. For example, therapeutic agents which are highly effective and display the fewest side effects may be selected for a patient based on his/her pharmacogenomic profile.

In another embodiment, SECP, fragments of SECP, or antibodies specific for SECP may be used as elements on a microarray. The microarray may be used to monitor or measure protein-protein interactions, drug-target interactions, and gene expression profiles, as described above.

A particular embodiment relates to the use of the polynucleotides of the present invention to generate a transcript image of a tissue or cell type. A transcript image represents the global pattern of gene expression by a particular tissue or cell type. Global gene expression patterns are analyzed by quantifying the number of expressed genes and their relative abundance under given conditions and at a given time. (See Seilhamer et al., "Comparative Gene Transcript Analysis," U.S. Patent Number 5,840,484, expressly incorporated by reference herein.) Thus a transcript image may be generated by hybridizing the polynucleotides of the present invention or their complements to the totality of transcripts or reverse transcripts of a particular tissue or cell type. In one embodiment, the hybridization takes place in high-throughput format, wherein the polynucleotides of the present invention or their complements comprise a subset of a plurality of elements on a microarray. The resultant transcript image would provide a profile of gene activity.

Transcript images may be generated using transcripts isolated from tissues, cell lines, biopsies, or other biological samples. The transcript image may thus reflect gene expression in vivo, as in the case of a tissue or biopsy sample, or in vitro, as in the case of a cell line.

Transcript images which profile the expression of the polynucleotides of the present invention may also be used in conjunction with in vitro model systems and preclinical evaluation of pharmaceuticals, as well as toxicological testing of industrial and naturally-occurring environmental compounds. All compounds induce characteristic gene expression patterns, frequently termed molecular fingerprints or toxicant signatures, which are indicative of mechanisms of action and toxicity (Nuwaysir, E.F. et al. (1999) Mol. Carcinog. 24:153-159; Steiner, S. and N.L. Anderson (2000) Toxicol. Lett. 112-113:467-471, expressly incorporated by reference herein). If a test compound has a signature similar to that of a compound with known toxicity, it is likely to share those toxic properties. These fingerprints or signatures are most useful and refined when they contain expression information from a large number of genes and gene families. Ideally, a genome-wide measurement of expression provides the highest quality signature. Even genes whose expression is not altered by any tested compounds are important as well, as the levels of expression of these genes are used to normalize the rest of the expression data. The normalization procedure is useful for comparison of expression data after treatment with different compounds. While the assignment of gene function to elements of a toxicant signature aids in interpretation of toxicity mechanisms, knowledge of gene function is not necessary for the statistical matching of signatures which leads to prediction of toxicity. (See, for example, Press Release 00-02 from the National Institute of Environmental Health Sciences, released February 29, 2000, available at http://www.niehs.nih.gov/oc/news/toxchip.htm.) Therefore, it is important and desirable in toxicological screening using toxicant signatures to include all expressed gene sequences.

In one embodiment, the toxicity of a test compound is assessed by treating a biological sample containing nucleic acids with the test compound. Nucleic acids that are expressed in the treated biological sample are hybridized with one or more probes specific to the polynucleotides of the present invention, so that transcript levels corresponding to the polynucleotides of the present invention may be quantified. The transcript levels in the treated biological sample are compared with levels in an untreated biological sample. Differences in the transcript levels between the two samples are indicative of a toxic response caused by the test compound in the treated sample.

Another particular embodiment relates to the use of the polypeptide sequences of the present invention to analyze the proteome of a tissue or cell type. The term proteome refers to the global pattern of protein expression in a particular tissue or cell type. Each protein component of a proteome can be subjected individually to further analysis. Proteome expression patterns, or profiles, are analyzed by quantifying the number of expressed proteins and their relative abundance under given conditions and at a given time. A profile of a cell's proteome may thus be generated by separating and analyzing the polypeptides of a particular tissue or cell type. In one embodiment, the separation is achieved using two-dimensional gel electrophoresis, in which proteins from a sample are

separated by isoelectric focusing in the first dimension, and then according to molecular weight by sodium dodecyl sulfate slab gel electrophoresis in the second dimension (Steiner and Anderson, supra). The proteins are visualized in the gel as discrete and uniquely positioned spots, typically by staining the gel with an agent such as Coomassie Blue or silver or fluorescent stains. The optical density of each protein spot is generally proportional to the level of the protein in the sample. The optical densities of equivalently positioned protein spots from different samples, for example, from biological samples either treated or untreated with a test compound or therapeutic agent, are compared to identify any changes in protein spot density related to the treatment. The proteins in the spots are partially sequenced using, for example, standard methods employing chemical or enzymatic cleavage followed by mass spectrometry. The identity of the protein in a spot may be determined by comparing its partial sequence, preferably of at least 5 contiguous amino acid residues, to the polypeptide sequences of the present invention. In some cases, further sequence data may be obtained for definitive protein identification.

A proteomic profile may also be generated using antibodies specific for SECP to quantify the levels of SECP expression. In one embodiment, the antibodies are used as elements on a microarray, and protein expression levels are quantified by exposing the microarray to the sample and detecting the levels of protein bound to each array element (Lueking, A. et al. (1999) Anal. Biochem. 270:103-111; Mendoze, L.G. et al. (1999) Biotechniques 27:778-788). Detection may be performed by a variety of methods known in the art, for example, by reacting the proteins in the sample with a thiol- or amino-reactive fluorescent compound and detecting the amount of fluorescence bound at each array element.

Toxicant signatures at the proteome level are also useful for toxicological screening, and should be analyzed in parallel with toxicant signatures at the transcript level. There is a poor correlation between transcript and protein abundances for some proteins in some tissues (Anderson, N.L. and J. Seilhamer (1997) Electrophoresis 18:533-537), so proteome toxicant signatures may be useful in the analysis of compounds which do not significantly affect the transcript image, but which alter the proteomic profile. In addition, the analysis of transcripts in body fluids is difficult, due to rapid degradation of mRNA, so proteomic profiling may be more reliable and informative in such cases.

In another embodiment, the toxicity of a test compound is assessed by treating a biological sample containing proteins with the test compound. Proteins that are expressed in the treated biological sample are separated so that the amount of each protein can be quantified. The amount of each protein is compared to the amount of the corresponding protein in an untreated biological sample. A difference in the amount of protein between the two samples is indicative of a toxic response to the test compound in the treated sample. Individual proteins are identified by sequencing

the amino acid residues of the individual proteins and comparing these partial sequences to the polypeptides of the present invention.

In another embodiment, the toxicity of a test compound is assessed by treating a biological sample containing proteins with the test compound. Proteins from the biological sample are incubated with antibodies specific to the polypeptides of the present invention. The amount of protein recognized by the antibodies is quantified. The amount of protein in the treated biological sample is compared with the amount in an untreated biological sample. A difference in the amount of protein between the two samples is indicative of a toxic response to the test compound in the treated sample.

Microarrays may be prepared, used, and analyzed using methods known in the art. (See, e.g., Brennan, T.M. et al. (1995) U.S. Patent No. 5,474,796; Schena, M. et al. (1996) Proc. Natl. Acad. Sci. USA 93:10614-10619; Baldeschweiler et al. (1995) PCT application WO95/251116; Shalon, D. et al. (1995) PCT application WO95/35505; Heller, R.A. et al. (1997) Proc. Natl. Acad. Sci. USA 94:2150-2155; and Heller, M.J. et al. (1997) U.S. Patent No. 5,605,662.) Various types of microarrays are well known and thoroughly described in DNA Microarrays: A Practical Approach, M. Schena, ed. (1999) Oxford University Press, London, hereby expressly incorporated by reference.

In another embodiment of the invention, nucleic acid sequences encoding SECP may be used to generate hybridization probes useful in mapping the naturally occurring genomic sequence. Either coding or noncoding sequences may be used, and in some instances, noncoding sequences may be preferable over coding sequences. For example, conservation of a coding sequence among members of a multi-gene family may potentially cause undesired cross hybridization during chromosomal mapping. The sequences may be mapped to a particular chromosome, to a specific region of a chromosome, or to artificial chromosome constructions, e.g., human artificial chromosomes (HACs), yeast artificial chromosomes (YACs), bacterial artificial chromosomes (BACs), bacterial P1 constructions, or single chromosome cDNA libraries. (See, e.g., Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355; Price, C.M. (1993) Blood Rev. 7:127-134; and Trask, B.J. (1991) Trends Genet. 7:149-154.) Once mapped, the nucleic acid sequences of the invention may be used to develop genetic linkage maps, for example, which correlate the inheritance of a disease state with the inheritance of a particular chromosome region or restriction fragment length polymorphism (RFLP). (See, for example, Lander, E.S. and D. Botstein (1986) Proc. Natl. Acad. Sci. USA 83:7353-7357.)

Fluorescent in situ hybridization (FISH) may be correlated with other physical and genetic map data. (See, e.g., Heinz-Ulrich, et al. (1995) in Meyers, supra, pp. 965-968.) Examples of genetic map data can be found in various scientific journals or at the Online Mendelian Inheritance in Man (OMIM) World Wide Web site. Correlation between the location of the gene encoding SECP on a physical map and a specific disorder, or a predisposition to a specific disorder, may help define the

region of DNA associated with that disorder and thus may further positional cloning efforts.

In situ hybridization of chromosomal preparations and physical mapping techniques, such as linkage analysis using established chromosomal markers, may be used for extending genetic maps. Often the placement of a gene on the chromosome of another mammalian species, such as mouse, may reveal associated markers even if the exact chromosomal locus is not known. This information is valuable to investigators searching for disease genes using positional cloning or other gene discovery techniques. Once the gene or genes responsible for a disease or syndrome have been crudely localized by genetic linkage to a particular genomic region, e.g., ataxia-telangiectasia to 11q22-23, any sequences mapping to that area may represent associated or regulatory genes for further investigation. (See, e.g., Gatti, R.A. et al. (1988) Nature 336:577-580.) The nucleotide sequence of the instant invention may also be used to detect differences in the chromosomal location due to translocation, inversion, etc., among normal, carrier, or affected individuals.

In another embodiment of the invention, SECP, its catalytic or immunogenic fragments, or oligopeptides thereof can be used for screening libraries of compounds in any of a variety of drug screening techniques. The fragment employed in such screening may be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. The formation of binding complexes between SECP and the agent being tested may be measured.

Another technique for drug screening provides for high throughput screening of compounds having suitable binding affinity to the protein of interest. (See, e.g., Geysen, et al. (1984) PCT application WO84/03564.) In this method, large numbers of different small test compounds are synthesized on a solid substrate. The test compounds are reacted with SECP, or fragments thereof, and washed. Bound SECP is then detected by methods well known in the art. Purified SECP can also be coated directly onto plates for use in the aforementioned drug screening techniques. Alternatively, non-neutralizing antibodies can be used to capture the peptide and immobilize it on a solid support.

In another embodiment, one may use competitive drug screening assays in which neutralizing antibodies capable of binding SECP specifically compete with a test compound for binding SECP. In this manner, antibodies can be used to detect the presence of any peptide which shares one or more antigenic determinants with SECP.

In additional embodiments, the nucleotide sequences which encode SECP may be used in any molecular biology techniques that have yet to be developed, provided the new techniques rely on properties of nucleotide sequences that are currently known, including, but not limited to, such properties as the triplet genetic code and specific base pair interactions.

Without further elaboration, it is believed that one skilled in the art can, using the preceding description, utilize the present invention to its fullest extent. The following preferred specific

embodiments are, therefore, to be construed as merely illustrative, and not limitative of the remainder of the disclosure in any way whatsoever.

The disclosures of all patents, applications, and publications mentioned above and below, in particular U.S. Ser. No.60/214,601, U.S. Ser. No. 60/212,890, U.S. Ser. No. 60/222,372, U.S. Ser. No. 60/213,466, U.S. Ser. No. 60/231,435, and U.S. Ser. No. 60/232,889, are hereby expressly incorporated by reference.

EXAMPLES

I. Construction of cDNA Libraries

Incyte cDNAs were derived from cDNA libraries described in the LIFESEQ GOLD database (Incyte Genomics, Palo Alto CA) and shown in Table 4, column 5. Some tissues were homogenized and lysed in guanidinium isothiocyanate, while others were homogenized and lysed in phenol or in a suitable mixture of denaturants, such as TRIZOL (Life Technologies), a monophasic solution of phenol and guanidine isothiocyanate. The resulting lysates were centrifuged over CsCl cushions or extracted with chloroform. RNA was precipitated from the lysates with either isopropanol or sodium acetate and ethanol, or by other routine methods.

Phenol extraction and precipitation of RNA were repeated as necessary to increase RNA purity. In some cases, RNA was treated with DNase. For most libraries, poly(A)+ RNA was isolated using oligo d(T)-coupled paramagnetic particles (Promega), OLIGOTEX latex particles (QIAGEN, Chatsworth CA), or an OLIGOTEX mRNA purification kit (QIAGEN). Alternatively, RNA was isolated directly from tissue lysates using other RNA isolation kits, e.g., the POLY(A)PURE mRNA purification kit (Ambion, Austin TX).

In some cases, Stratagene was provided with RNA and constructed the corresponding cDNA libraries. Otherwise, cDNA was synthesized and cDNA libraries were constructed with the UNIZAP vector system (Stratagene) or SUPERSCRIPT plasmid system (Life Technologies), using the recommended procedures or similar methods known in the art. (See, e.g., Ausubel, 1997, supra, units 5.1-6.6.) Reverse transcription was initiated using oligo d(T) or random primers. Synthetic oligonucleotide adapters were ligated to double stranded cDNA, and the cDNA was digested with the appropriate restriction enzyme or enzymes. For most libraries, the cDNA was size-selected (300-1000 bp) using SEPHACRYL S1000, SEPHAROSE CL2B, or SEPHAROSE CL4B column chromatography (Amersham Pharmacia Biotech) or preparative agarose gel electrophoresis. cDNAs were ligated into compatible restriction enzyme sites of the polylinker of a suitable plasmid, e.g., PBLUESCRIPT plasmid (Stratagene), PSPORT1 plasmid (Life Technologies), PCDNA2.1 plasmid (Invitrogen, Carlsbad CA), PBK-CMV plasmid (Stratagene), or pINCY (Incyte Genomics, Palo Alto CA), or derivatives thereof. Recombinant plasmids were transformed into competent E. coli cells

including XL1-Blue, XL1-BlueMRF, or SOLR from Stratagene or DH5α, DH10B, or ElectroMAX DH10B from Life Technologies.

II. Isolation of cDNA Clones

Plasmids obtained as described in Example I were recovered from host cells by in vivo excision using the UNIZAP vector system (Stratagene) or by cell lysis. Plasmids were purified using at least one of the following: a Magic or WIZARD Minipreps DNA purification system (Promega); an AGTC Miniprep purification kit (Edge Biosystems, Gaithersburg MD); and QIAWELL 8 Plasmid, QIAWELL 8 Plus Plasmid, QIAWELL 8 Ultra Plasmid purification systems or the R.E.A.L. PREP 96 plasmid purification kit from QIAGEN. Following precipitation, plasmids were resuspended in 0.1 ml of distilled water and stored, with or without lyophilization, at 4°C.

Alternatively, plasmid DNA was amplified from host cell lysates using direct link PCR in a high-throughput format (Rao, V.B. (1994) Anal. Biochem. 216:1-14). Host cell lysis and thermal cycling steps were carried out in a single reaction mixture. Samples were processed and stored in 384-well plates, and the concentration of amplified plasmid DNA was quantified fluorometrically using PICOGREEN dye (Molecular Probes, Eugene OR) and a FLUOROSKAN II fluorescence scanner (Labsystems Oy, Helsinki, Finland).

III. Sequencing and Analysis

Incyte cDNA recovered in plasmids as described in Example II were sequenced as follows. Sequencing reactions were processed using standard methods or high-throughput instrumentation such as the ABI CATALYST 800 (Applied Biosystems) thermal cycler or the PTC-200 thermal cycler (MJ Research) in conjunction with the HYDRA microdispenser (Robbins Scientific) or the MICROLAB 2200 (Hamilton) liquid transfer system. cDNA sequencing reactions were prepared using reagents provided by Amersham Pharmacia Biotech or supplied in ABI sequencing kits such as the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Applied Biosystems). Electrophoretic separation of cDNA sequencing reactions and detection of labeled polynucleotides were carried out using the MEGABACE 1000 DNA sequencing system (Molecular Dynamics); the ABI PRISM 373 or 377 sequencing system (Applied Biosystems) in conjunction with standard ABI protocols and base calling software; or other sequence analysis systems known in the art. Reading frames within the cDNA sequences were identified using standard methods (reviewed in Ausubel, 1997, supra, unit 7.7). Some of the cDNA sequences were selected for extension using the techniques disclosed in Example VIII.

The polynucleotide sequences derived from Incyte cDNAs were validated by removing vector, linker, and poly(A) sequences and by masking ambiguous bases, using algorithms and programs based on BLAST, dynamic programming, and dinucleotide nearest neighbor analysis. The Incyte cDNA sequences or translations thereof were then queried against a selection of public

databases such as the GenBank primate, rodent, mammalian, vertebrate, and eukaryote databases, and BLOCKS, PRINTS, DOMO, PRODOM, and hidden Markov model (HMM)-based protein family databases such as PFAM. (HMM is a probabilistic approach which analyzes consensus primary structures of gene families. See, for example, Eddy, S.R. (1996) Curr. Opin. Struct. Biol. 6:361-365.) The queries were performed using programs based on BLAST, FASTA, BLIMPS, and HMMER. The Incyte cDNA sequences were assembled to produce full length polynucleotide sequences. Alternatively, GenBank cDNAs, GenBank ESTs, stitched sequences, stretched sequences, or Genscan-predicted coding sequences (see Examples IV and V) were used to extend Incyte cDNA assemblages to full length. Assembly was performed using programs based on Phred, Phrap, and Consed, and cDNA assemblages were screened for open reading frames using programs based on GeneMark, BLAST, and FASTA. The full length polynucleotide sequences were translated to derive the corresponding full length polypeptide sequences. Alternatively, a polypeptide of the invention may begin at any of the methionine residues of the full length translated polypeptide. Full length polypeptide sequences were subsequently analyzed by querying against databases such as the GenBank protein databases (genpept), SwissProt, BLOCKS, PRINTS, DOMO, PRODOM, Prosite, and hidden Markov model (HMM)-based protein family databases such as PFAM. Full length polynucleotide sequences are also analyzed using MACDNASIS PRO software (Hitachi Software Engineering, South San Francisco CA) and LASERGENE software (DNASTAR). Polynucleotide and polypeptide sequence alignments are generated using default parameters specified by the CLUSTAL algorithm as incorporated into the MEGALIGN multisequence alignment program (DNASTAR), which also calculates the percent identity between aligned sequences.

Table 7 summarizes the tools, programs, and algorithms used for the analysis and assembly of Incyte cDNA and full length sequences and provides applicable descriptions, references, and threshold parameters. The first column of Table 7 shows the tools, programs, and algorithms used, the second column provides brief descriptions thereof, the third column presents appropriate references, all of which are incorporated by reference herein in their entirety, and the fourth column presents, where applicable, the scores, probability values, and other parameters used to evaluate the strength of a match between two sequences (the higher the score or the lower the probability value, the greater the identity between two sequences).

The programs described above for the assembly and analysis of full length polynucleotide and polypeptide sequences were also used to identify polynucleotide sequence fragments from SEQ ID NO:45-88. Fragments from about 20 to about 4000 nucleotides which are useful in hybridization and amplification technologies are described in Table 4, column 4.

IV. Identification and Editing of Coding Sequences from Genomic DNA

Putative secreted proteins were initially identified by running the Genscan gene identification

program against public genomic sequence databases (e.g., gbpri and gbhtg). Genscan is a general-purpose gene identification program which analyzes genomic DNA sequences from a variety of organisms (See Burge, C. and S. Karlin (1997) J. Mol. Biol. 268:78-94, and Burge, C. and S. Karlin (1998) Curr. Opin. Struct. Biol. 8:346-354). The program concatenates predicted exons to form an assembled cDNA sequence extending from a methionine to a stop codon. The output of Genscan is a FASTA database of polynucleotide and polypeptide sequences. The maximum range of sequence for Genscan to analyze at once was set to 30 kb. To determine which of these Genscan predicted cDNA sequences encode secreted proteins, the encoded polypeptides were analyzed by querying against PFAM models for secreted proteins. Potential secreted proteins were also identified by homology to Incyte cDNA sequences that had been annotated as secreted proteins. These selected Genscan-predicted sequences were then compared by BLAST analysis to the genpept and gbpri public databases. Where necessary, the Genscan-predicted sequences were then edited by comparison to the top BLAST hit from genpept to correct errors in the sequence predicted by Genscan, such as extra or omitted exons. BLAST analysis was also used to find any Incyte cDNA or public cDNA coverage of the Genscan-predicted sequences, thus providing evidence for transcription. When Incyte cDNA coverage was available, this information was used to correct or confirm the Genscan predicted sequence. Full length polynucleotide sequences were obtained by assembling Genscan-predicted coding sequences with Incyte cDNA sequences and/or public cDNA sequences using the assembly process described in Example III. Alternatively, full length polynucleotide sequences were derived entirely from edited or unedited Genscan-predicted coding sequences.

V. Assembly of Genomic Sequence Data with cDNA Sequence Data

"Stitched" Sequences

Partial cDNA sequences were extended with exons predicted by the Genscan gene identification program described in Example IV. Partial cDNAs assembled as described in Example III were mapped to genomic DNA and parsed into clusters containing related cDNAs and Genscan exon predictions from one or more genomic sequences. Each cluster was analyzed using an algorithm based on graph theory and dynamic programming to integrate cDNA and genomic information, generating possible splice variants that were subsequently confirmed, edited, or extended to create a full length sequence. Sequence intervals in which the entire length of the interval was present on more than one sequence in the cluster were identified, and intervals thus identified were considered to be equivalent by transitivity. For example, if an interval was present on a cDNA and two genomic sequences, then all three intervals were considered to be equivalent. This process allows unrelated but consecutive genomic sequences to be brought together, bridged by cDNA sequence. Intervals thus identified were then "stitched" together by the stitching algorithm in the order that they appear along their parent sequences to generate the longest possible sequence, as well as sequence variants.

Linkages between intervals which proceed along one type of parent sequence (cDNA to cDNA or genomic sequence to genomic sequence) were given preference over linkages which change parent type (cDNA to genomic sequence). The resultant stitched sequences were translated and compared by BLAST analysis to the genpept and gbpri public databases. Incorrect exons predicted by Genscan were corrected by comparison to the top BLAST hit from genpept. Sequences were further extended with additional cDNA sequences, or by inspection of genomic DNA, when necessary.

"Stretched" Sequences

Partial DNA sequences were extended to full length with an algorithm based on BLAST analysis. First, partial cDNAs assembled as described in Example III were queried against public databases such as the GenBank primate, rodent, mammalian, vertebrate, and eukaryote databases using the BLAST program. The nearest GenBank protein homolog was then compared by BLAST analysis to either Incyte cDNA sequences or GenScan exon predicted sequences described in Example IV. A chimeric protein was generated by using the resultant high-scoring segment pairs (HSPs) to map the translated sequences onto the GenBank protein homolog. Insertions or deletions may occur in the chimeric protein with respect to the original GenBank protein homolog. The GenBank protein homolog, the chimeric protein, or both were used as probes to search for homologous genomic sequences from the public human genome databases. Partial DNA sequences were therefore "stretched" or extended by the addition of homologous genomic sequences. The resultant stretched sequences were examined to determine whether it contained a complete gene.

VI. Chromosomal Mapping of SECP Encoding Polynucleotides

The sequences which were used to assemble SEQ ID NO:45-88 were compared with sequences from the Incyte LIFESEQ database and public domain databases using BLAST and other implementations of the Smith-Waterman algorithm. Sequences from these databases that matched SEQ ID NO:45-88 were assembled into clusters of contiguous and overlapping sequences using assembly algorithms such as Phrap (Table 7). Radiation hybrid and genetic mapping data available from public resources such as the Stanford Human Genome Center (SHGC), Whitehead Institute for Genome Research (WIGR), and Généthon were used to determine if any of the clustered sequences had been previously mapped. Inclusion of a mapped sequence in a cluster resulted in the assignment of all sequences of that cluster, including its particular SEQ ID NO:, to that map location.

Map locations are represented by ranges, or intervals, of human chromosomes. The map position of an interval, in centiMorgans, is measured relative to the terminus of the chromosome's p-arm. (The centiMorgan (cM) is a unit of measurement based on recombination frequencies between chromosomal markers. On average, 1 cM is roughly equivalent to 1 megabase (Mb) of DNA in humans, although this can vary widely due to hot and cold spots of recombination.) The cM distances are based on genetic markers mapped by Généthon which provide boundaries for radiation

hybrid markers whose sequences were included in each of the clusters. Human genome maps and other resources available to the public, such as the NCBI "GeneMap'99" World Wide Web site (http://www.ncbi.nlm.nih.gov/genemap/), can be employed to determine if previously identified disease genes map within or in proximity to the intervals indicated above.

In this manner, SEQ ID NO:48 was mapped to chromosome 15 within the interval from 72.3 to 77.4 centiMorgans.

In this manner, SEQ ID NO:54 was mapped to chromosome 20 within the interval from 6.20 to 9.40 centiMorgans. SEQ ID NO:61 was mapped to chromosome 22 within the interval from 0.00 to 19.50 centiMorgans.

In this manner, SEQ ID NO:82 was mapped to chromosome 22 within the interval from 0.0 to 19.5 centiMorgans. SEQ ID NO:85 was mapped to chromosome 12 within the interval from 84.7 to 92.5 centiMorgans and from 137.5 to 145.7 centiMorgans. More than one map location is reported for SEQ ID NO:85, indicating that sequences having different map locations were assembled into a single cluster. This situation occurs, for example, when sequences having strong similarity, but not complete identity, are assembled into a single cluster.

In this manner, SEQ ID NO:66 was mapped to chromosome 16 within the interval from 65.60 to 72.60 centiMorgans. In this manner, SEQ ID NO:67 was mapped to chromosome 11 within the interval from 59.50 to 65.00 centiMorgans. In this manner, SEQ ID NO:69 was mapped to chromosome 6 within the interval from 132.70 to 144.40 centiMorgans.

VII. Analysis of Polynucleotide Expression

Northern analysis is a laboratory technique used to detect the presence of a transcript of a gene and involves the hybridization of a labeled nucleotide sequence to a membrane on which RNAs from a particular cell type or tissue have been bound. (See, e.g., Sambrook, supra, ch. 7; Ausubel (1995) supra, ch. 4 and 16.)

Analogous computer techniques applying BLAST were used to search for identical or related molecules in cDNA databases such as GenBank or LIFESEQ (Incyte Genomics). This analysis is much faster than multiple membrane-based hybridizations. In addition, the sensitivity of the computer search can be modified to determine whether any particular match is categorized as exact or similar. The basis of the search is the product score, which is defined as:

$$\frac{\text{BLAST Score} \times \text{Percent Identity}}{5 \times \text{minimum}\{\text{length(Seq. 1)}, \text{length(Seq. 2)}\}}$$

The product score takes into account both the degree of similarity between two sequences and the length of the sequence match. The product score is a normalized value between 0 and 100, and is

calculated as follows: the BLAST score is multiplied by the percent nucleotide identity and the product is divided by (5 times the length of the shorter of the two sequences). The BLAST score is calculated by assigning a score of +5 for every base that matches in a high-scoring segment pair (HSP), and -4 for every mismatch. Two sequences may share more than one HSP (separated by gaps). If there is more than one HSP, then the pair with the highest BLAST score is used to calculate the product score. The product score represents a balance between fractional overlap and quality in a BLAST alignment. For example, a product score of 100 is produced only for 100% identity over the entire length of the shorter of the two sequences being compared. A product score of 70 is produced either by 100% identity and 70% overlap at one end, or by 88% identity and 100% overlap at the other. A product score of 50 is produced either by 100% identity and 50% overlap at one end, or 79% identity and 100% overlap.

Alternatively, polynucleotide sequences encoding SECP are analyzed with respect to the tissue sources from which they were derived. For example, some full length sequences are assembled, at least in part, with overlapping Incyte cDNA sequences (see Example III). Each cDNA sequence is derived from a cDNA library constructed from a human tissue. Each human tissue is classified into one of the following organ/tissue categories: cardiovascular system; connective tissue; digestive system; embryonic structures; endocrine system; exocrine glands; genitalia, female; genitalia, male; germ cells; hemic and immune system; liver; musculoskeletal system; nervous system; pancreas; respiratory system; sense organs; skin; stomatognathic system; unclassified/mixed; or urinary tract. The number of libraries in each category is counted and divided by the total number of libraries across all categories. Similarly, each human tissue is classified into one of the following disease/condition categories: cancer, cell line, developmental, inflammation, neurological, trauma, cardiovascular, pooled, and other, and the number of libraries in each category is counted and divided by the total number of libraries across all categories. The resulting percentages reflect the tissue- and disease-specific expression of cDNA encoding SECP. cDNA sequences and cDNA library/tissue information are found in the LIFESEQ GOLD database (Incyte Genomics, Palo Alto CA).

VIII. Extension of SECP Encoding Polynucleotides

Full length polynucleotide sequences were also produced by extension of an appropriate fragment of the full length molecule using oligonucleotide primers designed from this fragment. One primer was synthesized to initiate 5' extension of the known fragment, and the other primer was synthesized to initiate 3' extension of the known fragment. The initial primers were designed using OLIGO 4.06 software (National Biosciences), or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the target sequence at temperatures of about 68°C to about 72°C. Any stretch of nucleotides which would result in hairpin structures and primer-primer dimerizations was avoided.

Selected human cDNA libraries were used to extend the sequence. If more than one extension was necessary or desired, additional or nested sets of primers were designed.

High fidelity amplification was obtained by PCR using methods well known in the art. PCR was performed in 96-well plates using the PTC-200 thermal cycler (MJ Research, Inc.). The reaction mix contained DNA template, 200 nmol of each primer, reaction buffer containing Mg^{2+}, $(NH_4)_2SO_4$, and 2-mercaptoethanol, Taq DNA polymerase (Amersham Pharmacia Biotech), ELONGASE enzyme (Life Technologies), and Pfu DNA polymerase (Stratagene), with the following parameters for primer pair PCI A and PCI B: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C. In the alternative, the parameters for primer pair T7 and SK+ were as follows: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 57°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C.

The concentration of DNA in each well was determined by dispensing 100 µl PICOGREEN quantitation reagent (0.25% (v/v) PICOGREEN; Molecular Probes, Eugene OR) dissolved in 1X TE and 0.5 µl of undiluted PCR product into each well of an opaque fluorimeter plate (Corning Costar, Acton MA), allowing the DNA to bind to the reagent. The plate was scanned in a Fluoroskan II (Labsystems Oy, Helsinki, Finland) to measure the fluorescence of the sample and to quantify the concentration of DNA. A 5 μl to 10 μl aliquot of the reaction mixture was analyzed by electrophoresis on a 1 % agarose gel to determine which reactions were successful in extending the sequence.

The extended nucleotides were desalted and concentrated, transferred to 384-well plates, digested with CviJI cholera virus endonuclease (Molecular Biology Research, Madison WI), and sonicated or sheared prior to religation into pUC 18 vector (Amersham Pharmacia Biotech). For shotgun sequencing, the digested nucleotides were separated on low concentration (0.6 to 0.8%) agarose gels, fragments were excised, and agar digested with Agar ACE (Promega). Extended clones were religated using T4 ligase (New England Biolabs, Beverly MA) into pUC 18 vector (Amersham Pharmacia Biotech), treated with Pfu DNA polymerase (Stratagene) to fill-in restriction site overhangs, and transfected into competent E. coli cells. Transformed cells were selected on antibiotic-containing media, and individual colonies were picked and cultured overnight at 37°C in 384-well plates in LB/2x carb liquid media.

The cells were lysed, and DNA was amplified by PCR using Taq DNA polymerase (Amersham Pharmacia Biotech) and Pfu DNA polymerase (Stratagene) with the following parameters: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 72°C, 2 min; Step 5: steps 2, 3, and 4 repeated 29 times; Step 6: 72°C, 5 min; Step 7: storage at 4°C. DNA was quantified by PICOGREEN reagent (Molecular Probes) as described above. Samples with low DNA

recoveries were reamplified using the same conditions as described above. Samples were diluted with 20% dimethysulfoxide (1:2, v/v), and sequenced using DYENAMIC energy transfer sequencing primers and the DYENAMIC DIRECT kit (Amersham Pharmacia Biotech) or the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Applied Biosystems).

In like manner, full length polynucleotide sequences are verified using the above procedure or are used to obtain 5' regulatory sequences using the above procedure along with oligonucleotides designed for such extension, and an appropriate genomic library.

IX. Labeling and Use of Individual Hybridization Probes

Hybridization probes derived from SEQ ID NO:45-88 are employed to screen cDNAs, genomic DNAs, or mRNAs. Although the labeling of oligonucleotides, consisting of about 20 base pairs, is specifically described, essentially the same procedure is used with larger nucleotide fragments. Oligonucleotides are designed using state-of-the-art software such as OLIGO 4.06 software (National Biosciences) and labeled by combining 50 pmol of each oligomer, 250 μCi of [γ-^{32}P] adenosine triphosphate (Amersham Pharmacia Biotech), and T4 polynucleotide kinase (DuPont NEN, Boston MA). The labeled oligonucleotides are substantially purified using a SEPHADEX G-25 superfine size exclusion dextran bead column (Amersham Pharmacia Biotech). An aliquot containing 10^7 counts per minute of the labeled probe is used in a typical membrane-based hybridization analysis of human genomic DNA digested with one of the following endonucleases: Ase I, Bgl II, Eco RI, Pst I, Xba I, or Pvu II (DuPont NEN).

The DNA from each digest is fractionated on a 0.7% agarose gel and transferred to nylon membranes (Nytran Plus, Schleicher & Schuell, Durham NH). Hybridization is carried out for 16 hours at 40°C. To remove nonspecific signals, blots are sequentially washed at room temperature under conditions of up to, for example, 0.1 x saline sodium citrate and 0.5% sodium dodecyl sulfate. Hybridization patterns are visualized using autoradiography or an alternative imaging means and compared.

X. Microarrays

The linkage or synthesis of array elements upon a microarray can be achieved utilizing photolithography, piezoelectric printing (ink-jet printing, See, e.g., Baldeschweiler, supra.), mechanical microspotting technologies, and derivatives thereof. The substrate in each of the aforementioned technologies should be uniform and solid with a non-porous surface (Schena (1999), supra). Suggested substrates include silicon, silica, glass slides, glass chips, and silicon wafers. Alternatively, a procedure analogous to a dot or slot blot may also be used to arrange and link elements to the surface of a substrate using thermal, UV, chemical, or mechanical bonding procedures. A typical array may be produced using available methods and machines well known to those of ordinary skill in the art and may contain any appropriate number of elements. (See, e.g.,

Schena, M. et al. (1995) Science 270:467-470; Shalon, D. et al. (1996) Genome Res. 6:639-645; Marshall, A. and J. Hodgson (1998) Nat. Biotechnol. 16:27-31.)

Full length cDNAs, Expressed Sequence Tags (ESTs), or fragments or oligomers thereof may comprise the elements of the microarray. Fragments or oligomers suitable for hybridization can be selected using software well known in the art such as LASERGENE software (DNASTAR). The array elements are hybridized with polynucleotides in a biological sample. The polynucleotides in the biological sample are conjugated to a fluorescent label or other molecular tag for ease of detection. After hybridization, nonhybridized nucleotides from the biological sample are removed, and a fluorescence scanner is used to detect hybridization at each array element. Alternatively, laser desorbtion and mass spectrometry may be used for detection of hybridization. The degree of complementarity and the relative abundance of each polynucleotide which hybridizes to an element on the microarray may be assessed. In one embodiment, microarray preparation and usage is described in detail below.

Tissue or Cell Sample Preparation

Total RNA is isolated from tissue samples using the guanidinium thiocyanate method and poly(A)$^+$ RNA is purified using the oligo-(dT) cellulose method. Each poly(A)$^+$ RNA sample is reverse transcribed using MMLV reverse-transcriptase, 0.05 pg/µl oligo-(dT) primer (21mer), 1X first strand buffer, 0.03 units/µl RNase inhibitor, 500 µM dATP, 500 µM dGTP, 500 µM dTTP, 40 µM dCTP, 40 µM dCTP-Cy3 (BDS) or dCTP-Cy5 (Amersham Pharmacia Biotech). The reverse transcription reaction is performed in a 25 ml volume containing 200 ng poly(A)$^+$ RNA with GEMBRIGHT kits (Incyte). Specific control poly(A)$^+$ RNAs are synthesized by in vitro transcription from non-coding yeast genomic DNA. After incubation at 37°C for 2 hr, each reaction sample (one with Cy3 and another with Cy5 labeling) is treated with 2.5 ml of 0.5M sodium hydroxide and incubated for 20 minutes at 85°C to the stop the reaction and degrade the RNA. Samples are purified using two successive CHROMA SPIN 30 gel filtration spin columns (CLONTECH Laboratories, Inc. (CLONTECH), Palo Alto CA) and after combining, both reaction samples are ethanol precipitated using 1 ml of glycogen (1 mg/ml), 60 ml sodium acetate, and 300 ml of 100% ethanol. The sample is then dried to completion using a SpeedVAC (Savant Instruments Inc., Holbrook NY) and resuspended in 14 µl 5X SSC/0.2% SDS.

Microarray Preparation

Sequences of the present invention are used to generate array elements. Each array element is amplified from bacterial cells containing vectors with cloned cDNA inserts. PCR amplification uses primers complementary to the vector sequences flanking the cDNA insert. Array elements are amplified in thirty cycles of PCR from an initial quantity of 1-2 ng to a final quantity greater than 5 µg. Amplified array elements are then purified using SEPHACRYL-400 (Amersham Pharmacia

Biotech).

Purified array elements are immobilized on polymer-coated glass slides. Glass microscope slides (Corning) are cleaned by ultrasound in 0.1% SDS and acetone, with extensive distilled water washes between and after treatments. Glass slides are etched in 4% hydrofluoric acid (VWR Scientific Products Corporation (VWR), West Chester PA), washed extensively in distilled water, and coated with 0.05% aminopropyl silane (Sigma) in 95% ethanol. Coated slides are cured in a 110°C oven.

Array elements are applied to the coated glass substrate using a procedure described in US Patent No. 5,807,522, incorporated herein by reference. 1 µl of the array element DNA, at an average concentration of 100 ng/µl, is loaded into the open capillary printing element by a high-speed robotic apparatus. The apparatus then deposits about 5 nl of array element sample per slide.

Microarrays are UV-crosslinked using a STRATALINKER UV-crosslinker (Stratagene). Microarrays are washed at room temperature once in 0.2% SDS and three times in distilled water. Non-specific binding sites are blocked by incubation of microarrays in 0.2% casein in phosphate buffered saline (PBS) (Tropix, Inc., Bedford MA) for 30 minutes at 60°C followed by washes in 0.2% SDS and distilled water as before.

Hybridization

Hybridization reactions contain 9 µl of sample mixture consisting of 0.2 µg each of Cy3 and Cy5 labeled cDNA synthesis products in 5X SSC, 0.2% SDS hybridization buffer. The sample mixture is heated to 65°C for 5 minutes and is aliquoted onto the microarray surface and covered with an 1.8 cm^2 coverslip. The arrays are transferred to a waterproof chamber having a cavity just slightly larger than a microscope slide. The chamber is kept at 100% humidity internally by the addition of 140 µl of 5X SSC in a corner of the chamber. The chamber containing the arrays is incubated for about 6.5 hours at 60°C. The arrays are washed for 10 min at 45°C in a first wash buffer (1X SSC, 0.1% SDS), three times for 10 minutes each at 45°C in a second wash buffer (0.1X SSC), and dried.

Detection

Reporter-labeled hybridization complexes are detected with a microscope equipped with an Innova 70 mixed gas 10 W laser (Coherent, Inc., Santa Clara CA) capable of generating spectral lines at 488 nm for excitation of Cy3 and at 632 nm for excitation of Cy5. The excitation laser light is focused on the array using a 20X microscope objective (Nikon, Inc., Melville NY). The slide containing the array is placed on a computer-controlled X-Y stage on the microscope and raster-scanned past the objective. The 1.8 cm x 1.8 cm array used in the present example is scanned with a resolution of 20 micrometers.

In two separate scans, a mixed gas multiline laser excites the two fluorophores sequentially. Emitted light is split, based on wavelength, into two photomultiplier tube detectors (PMT R1477, Hamamatsu Photonics Systems, Bridgewater NJ) corresponding to the two fluorophores. Appropriate filters positioned between the array and the photomultiplier tubes are used to filter the signals. The emission maxima of the fluorophores used are 565 nm for Cy3 and 650 nm for Cy5. Each array is typically scanned twice, one scan per fluorophore using the appropriate filters at the laser source, although the apparatus is capable of recording the spectra from both fluorophores simultaneously.

The sensitivity of the scans is typically calibrated using the signal intensity generated by a cDNA control species added to the sample mixture at a known concentration. A specific location on the array contains a complementary DNA sequence, allowing the intensity of the signal at that location to be correlated with a weight ratio of hybridizing species of 1:100,000. When two samples from different sources (e.g., representing test and control cells), each labeled with a different fluorophore, are hybridized to a single array for the purpose of identifying genes that are differentially expressed, the calibration is done by labeling samples of the calibrating cDNA with the two fluorophores and adding identical amounts of each to the hybridization mixture.

The output of the photomultiplier tube is digitized using a 12-bit RTI-835H analog-to-digital (A/D) conversion board (Analog Devices, Inc., Norwood MA) installed in an IBM-compatible PC computer. The digitized data are displayed as an image where the signal intensity is mapped using a linear 20-color transformation to a pseudocolor scale ranging from blue (low signal) to red (high signal). The data is also analyzed quantitatively. Where two different fluorophores are excited and measured simultaneously, the data are first corrected for optical crosstalk (due to overlapping emission spectra) between the fluorophores using each fluorophore's emission spectrum.

A grid is superimposed over the fluorescence signal image such that the signal from each spot is centered in each element of the grid. The fluorescence signal within each element is then integrated to obtain a numerical value corresponding to the average intensity of the signal. The software used for signal analysis is the GEMTOOLS gene expression analysis program (Incyte).

XI. Complementary Polynucleotides

Sequences complementary to the SECP-encoding sequences, or any parts thereof, are used to detect, decrease, or inhibit expression of naturally occurring SECP. Although use of oligonucleotides comprising from about 15 to 30 base pairs is described, essentially the same procedure is used with smaller or with larger sequence fragments. Appropriate oligonucleotides are designed using OLIGO 4.06 software (National Biosciences) and the coding sequence of SECP. To inhibit transcription, a complementary oligonucleotide is designed from the most unique 5' sequence and used to prevent promoter binding to the coding sequence. To inhibit translation, a complementary oligonucleotide is designed to prevent ribosomal binding to the SECP-encoding transcript.

XII. Expression of SECP

Expression and purification of SECP is achieved using bacterial or virus-based expression systems. For expression of SECP in bacteria, cDNA is subcloned into an appropriate vector containing an antibiotic resistance gene and an inducible promoter that directs high levels of cDNA transcription. Examples of such promoters include, but are not limited to, the *trp-lac* (*tac*) hybrid promoter and the T5 or T7 bacteriophage promoter in conjunction with the *lac* operator regulatory element. Recombinant vectors are transformed into suitable bacterial hosts, e.g., BL21(DE3). Antibiotic resistant bacteria express SECP upon induction with isopropyl beta-D-thiogalactopyranoside (IPTG). Expression of SECP in eukaryotic cells is achieved by infecting insect or mammalian cell lines with recombinant Autographica californica nuclear polyhedrosis virus (AcMNPV), commonly known as baculovirus. The nonessential polyhedrin gene of baculovirus is replaced with cDNA encoding SECP by either homologous recombination or bacterial-mediated transposition involving transfer plasmid intermediates. Viral infectivity is maintained and the strong polyhedrin promoter drives high levels of cDNA transcription. Recombinant baculovirus is used to infect Spodoptera frugiperda (Sf9) insect cells in most cases, or human hepatocytes, in some cases. Infection of the latter requires additional genetic modifications to baculovirus. (See Engelhard, E.K. et al. (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. et al. (1996) Hum. Gene Ther. 7:1937-1945.)

In most expression systems, SECP is synthesized as a fusion protein with, e.g., glutathione S-transferase (GST) or a peptide epitope tag, such as FLAG or 6-His, permitting rapid, single-step, affinity-based purification of recombinant fusion protein from crude cell lysates. GST, a 26-kilodalton enzyme from Schistosoma japonicum, enables the purification of fusion proteins on immobilized glutathione under conditions that maintain protein activity and antigenicity (Amersham Pharmacia Biotech). Following purification, the GST moiety can be proteolytically cleaved from SECP at specifically engineered sites. FLAG, an 8-amino acid peptide, enables immunoaffinity purification using commercially available monoclonal and polyclonal anti-FLAG antibodies (Eastman Kodak). 6-His, a stretch of six consecutive histidine residues, enables purification on metal-chelate resins (QIAGEN). Methods for protein expression and purification are discussed in Ausubel (1995, supra, ch. 10 and 16). Purified SECP obtained by these methods can be used directly in the assays shown in Examples XVI and XVII, where applicable.

XIII. Functional Assays

SECP function is assessed by expressing the sequences encoding SECP at physiologically elevated levels in mammalian cell culture systems. cDNA is subcloned into a mammalian expression vector containing a strong promoter that drives high levels of cDNA expression. Vectors of choice include PCMV SPORT (Life Technologies) and PCR3.1 (Invitrogen, Carlsbad CA), both of which

contain the cytomegalovirus promoter. 5-10 μg of recombinant vector are transiently transfected into a human cell line, for example, an endothelial or hematopoietic cell line, using either liposome formulations or electroporation. 1-2 μg of an additional plasmid containing sequences encoding a marker protein are co-transfected. Expression of a marker protein provides a means to distinguish transfected cells from nontransfected cells and is a reliable predictor of cDNA expression from the recombinant vector. Marker proteins of choice include, e.g., Green Fluorescent Protein (GFP; Clontech), CD64, or a CD64-GFP fusion protein. Flow cytometry (FCM), an automated, laser optics-based technique, is used to identify transfected cells expressing GFP or CD64-GFP and to evaluate the apoptotic state of the cells and other cellular properties. FCM detects and quantifies the uptake of fluorescent molecules that diagnose events preceding or coincident with cell death. These events include changes in nuclear DNA content as measured by staining of DNA with propidium iodide; changes in cell size and granularity as measured by forward light scatter and 90 degree side light scatter; down-regulation of DNA synthesis as measured by decrease in bromodeoxyuridine uptake; alterations in expression of cell surface and intracellular proteins as measured by reactivity with specific antibodies; and alterations in plasma membrane composition as measured by the binding of fluorescein-conjugated Annexin V protein to the cell surface. Methods in flow cytometry are discussed in Ormerod, M.G. (1994) Flow Cytometry, Oxford, New York NY.

The influence of SECP on gene expression can be assessed using highly purified populations of cells transfected with sequences encoding SECP and either CD64 or CD64-GFP. CD64 and CD64-GFP are expressed on the surface of transfected cells and bind to conserved regions of human immunoglobulin G (IgG). Transfected cells are efficiently separated from nontransfected cells using magnetic beads coated with either human IgG or antibody against CD64 (DYNAL, Lake Success NY). mRNA can be purified from the cells using methods well known by those of skill in the art. Expression of mRNA encoding SECP and other genes of interest can be analyzed by northern analysis or microarray techniques.

XIV. Production of SECP Specific Antibodies

SECP substantially purified using polyacrylamide gel electrophoresis (PAGE; see, e.g., Harrington, M.G. (1990) Methods Enzymol. 182:488-495), or other purification techniques, is used to immunize rabbits and to produce antibodies using standard protocols.

Alternatively, the SECP amino acid sequence is analyzed using LASERGENE software (DNASTAR) to determine regions of high immunogenicity, and a corresponding oligopeptide is synthesized and used to raise antibodies by means known to those of skill in the art. Methods for selection of appropriate epitopes, such as those near the C-terminus or in hydrophilic regions are well described in the art. (See, e.g., Ausubel, 1995, supra, ch. 11.)

Typically, oligopeptides of about 15 residues in length are synthesized using an ABI 431A

peptide synthesizer (Applied Biosystems) using FMOC chemistry and coupled to KLH (Sigma-Aldrich, St. Louis MO) by reaction with N-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS) to increase immunogenicity. (See, e.g., Ausubel, 1995, supra.) Rabbits are immunized with the oligopeptide-KLH complex in complete Freund's adjuvant. Resulting antisera are tested for antipeptide and anti-SECP activity by, for example, binding the peptide or SECP to a substrate, blocking with 1% BSA, reacting with rabbit antisera, washing, and reacting with radio-iodinated goat anti-rabbit IgG.

XV. Purification of Naturally Occurring SECP Using Specific Antibodies

Naturally occurring or recombinant SECP is substantially purified by immunoaffinity chromatography using antibodies specific for SECP. An immunoaffinity column is constructed by covalently coupling anti-SECP antibody to an activated chromatographic resin, such as CNBr-activated SEPHAROSE (Amersham Pharmacia Biotech). After the coupling, the resin is blocked and washed according to the manufacturer's instructions.

Media containing SECP are passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of SECP (e.g., high ionic strength buffers in the presence of detergent). The column is eluted under conditions that disrupt antibody/SECP binding (e.g., a buffer of pH 2 to pH 3, or a high concentration of a chaotrope, such as urea or thiocyanate ion), and SECP is collected.

XVI. Identification of Molecules Which Interact with SECP

SECP, or biologically active fragments thereof, are labeled with ^{125}I Bolton-Hunter reagent. (See, e.g., Bolton A.E. and W.M. Hunter (1973) Biochem. J. 133:529-539.) Candidate molecules previously arrayed in the wells of a multi-well plate are incubated with the labeled SECP, washed, and any wells with labeled SECP complex are assayed. Data obtained using different concentrations of SECP are used to calculate values for the number, affinity, and association of SECP with the candidate molecules.

Alternatively, molecules interacting with SECP are analyzed using the yeast two-hybrid system as described in Fields, S. and O. Song (1989) Nature 340:245-246, or using commercially available kits based on the two-hybrid system, such as the MATCHMAKER system (Clontech).

SECP may also be used in the PATHCALLING process (CuraGen Corp., New Haven CT) which employs the yeast two-hybrid system in a high-throughput manner to determine all interactions between the proteins encoded by two large libraries of genes (Nandabalan, K. et al. (2000) U.S. Patent No. 6,057,101).

XVII. Demonstration of SECP Activity

An assay for the determination of SECP activity consists of an enzyme reaction mixture consisting of 25 mM Tris-HCl (pH 7.4), 0.25% Triton X-100, 5 MM $MnCl_2$, 5 mM CDP-choline, 5

mM 2-mercaptoethanol, 0.05 mM UDP-[^{14}C]GalNAc (4,000 cpm/nmol), 250 μM peptide, and varying amounts of SECP in a final volume of 100 μl. The reaction mixture is incubated for 10 min. at 37°C followed by Dowex 1 ion exchange (formic acid form) chromatography. Eluted peptide-containing fractions are subjected to scintillation counting. The amount of [^{14}C]GalNAc present in the peptide-containing fractions is proportional to SECP activity. Confirmation of substrate and SECP source can be evaluated by C-18 chromatography (C2C18 3.2 Smart System, Pharmacia Biotech Inc.) to ensure peptide stability and that incorporated [^{14}C]GalNAc is associated with the peptide (Sørensen,T. et al. (1995) J. Biol. Chem. 270:24166-24173).

Alternatively, an assay for growth stimulating or inhibiting activity of SECP measures the amount of DNA synthesis in Swiss mouse 3T3 cells (McKay, I. and Leigh, I., eds. (1993) Growth Factors: A Practical Approach, Oxford University Press, New York, NY). In this assay, varying amounts of SECP are added to quiescent 3T3 cultured cells in the presence of [^{3}H]thymidine, a radioactive DNA precursor. SECP for this assay can be obtained by recombinant means or from biochemical preparations. Incorporation of [^{3}H]thymidine into acid-precipitable DNA is measured over an appropriate time interval, and the amount incorporated is directly proportional to the amount of newly synthesized DNA. A linear dose-response curve over at least a hundred-fold SECP concentration range is indicative of growth modulating activity. One unit of activity per milliliter is defined as the concentration of SECP producing a 50% response level, where 100% represents maximal incorporation of [^{3}H]thymidine into acid-precipitable DNA.

Alternatively, an assay for SECP activity measures the stimulation or inhibition of neurotransmission in cultured cells. Cultured CHO fibroblasts are exposed to SECP. Following endocytic uptake of SECP, the cells are washed with fresh culture medium, and a whole cell voltage-clamped Xenopus myocyte is manipulated into contact with one of the fibroblasts in SECP-free medium. Membrane currents are recorded from the myocyte. Increased or decreased current relative to control values are indicative of neuromodulatory effects of SECP (Morimoto, T. et al. (1995) Neuron 15:689-696).

Alternatively, an assay for SECP activity measures the amount of SECP in secretory, membrane-bound organelles. Transfected cells as described above are harvested and lysed. The lysate is fractionated using methods known to those of skill in the art, for example, sucrose gradient ultracentrifugation. Such methods allow the isolation of subcellular components such as the Golgi apparatus, ER, small membrane-bound vesicles, and other secretory organelles. Immunoprecipitations from fractionated and total cell lysates are performed using SECP-specific antibodies, and immunoprecipitated samples are analyzed using SDS-PAGE and immunoblotting techniques. The concentration of SECP in secretory organelles relative to SECP in total cell lysate is proportional to the amount of SECP in transit through the secretory pathway.

In another alternative, SECP recognizes and precipitates antigen from serum. This activity can be measured by the quantitative precipitin reaction. (Golub, E. S. et al. (1987) Immunology: A Synthesis, Sinauer Associates, Sunderland, MA, pages 113-115.) SECP is isotopically labeled using methods known in the art. Various serum concentrations are added to constant amounts of labeled SECP. SECP-antigen complexes precipitate out of solution and are collected by centrifugation. The amount of precipitable SECP-antigen complex is proportional to the amount of radioisotope detected in the precipitate. The amount of precipitable SECP-antigen complex is plotted against the serum concentration. For various serum concentrations, a characteristic precipitation curve is obtained, in which the amount of precipitable SECP-antigen complex initially increases proportionately with increasing serum concentration, peaks at the equivalence point, and then decreases proportionately with further increases in serum concentration. Thus, the amount of precipitable SECP-antigen complex is a measure of SECP activity which is characterized by sensitivity to both limiting and excess quantities of antigen.

Alternatively, an assay for SECP activity measures the expression of SECP on the cell surface. cDNA encoding SECP is transfected into a non-leukocytic cell line. Cell surface proteins are labeled with biotin (de la Fuente, M.A. et.al. (1997) Blood 90:2398-2405). Immunoprecipitations are performed using SECP-specific antibodies, and immunoprecipitated samples are analyzed using SDS-PAGE and immunoblotting techniques. The ratio of labeled immunoprecipitant to unlabeled immunoprecipitant is proportional to the amount of SECP expressed on the cell surface.

Various modifications and variations of the described methods and systems of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with certain embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the invention which are obvious to those skilled in molecular biology or related fields are intended to be within the scope of the following claims.

Table 1

Incyte Project ID	Polypeptide SEQ ID NO:	Incyte Polypeptide ID	Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID
2101688	1	2101688CD1	45	2101688CB1
5452330	2	5452330CD1	46	5452330CB1
4362432	3	4362432CD1	47	4362432CB1
5308104	4	5308104CD1	48	5308104CB1
3092736	5	3092736CD1	49	3092736CB1
3580257	6	3580257CD1	50	3580257CB1
3634758	7	3634758CD1	51	3634758CB1
4027923	8	4027923CD1	52	4027923CB1
4348533	9	4348533CD1	53	4348533CB1
4521857	10	4521857CD1	54	4521857CB1
4722253	11	4722253CD1	55	4722253CB1
4878134	12	4878134CD1	56	4878134CB1
5050133	13	5050133CD1	57	5050133CB1
5630124	14	5630124CD1	58	5630124CB1
5677286	15	5677286CD1	59	5677286CB1
6436791	16	6436791CD1	60	6436791CB1
1820972	17	1820972CD1	61	1820972CB1
3286805	18	3286805CD1	62	3286805CB1
3506590	19	3506590CD1	63	3506590CB1
003600	20	003600CD1	64	003600CB1
1251534	21	1251534CD1	65	1251534CB1
1402211	22	1402211CD1	66	1402211CB1
1623474	23	1623474CD1	67	1623474CB1
1706443	24	1706443CD1	68	1706443CB1
1748627	25	1748627CD1	69	1748627CB1
1818332	26	1818332CD1	70	1818332CB1
1822832	27	1822832CD1	71	1822832CB1
1832219	28	1832219CD1	72	1832219CB1
1899010	29	1899010CD1	73	1899010CB1
2008768	30	2008768CD1	74	2008768CB1
2070984	31	2070984CD1	75	2070984CB1
2193240	32	2193240CD1	76	2193240CB1
2235177	33	2235177CD1	77	2235177CB1
2416227	34	2416227CD1	78	2416227CB1
2461076	35	2461076CD1	79	2461076CB1
1957517	36	1957517CD1	80	1957517CB1
866038	37	866038CD1	81	866038CB1
3869704	38	3869704CD1	82	3869704CB1
1415179	39	1415179CD1	83	1415179CB1

Table 1 (cont.)

Incyte Project ID	Polypeptide SEQ ID NO:	Incyte Polypeptide ID	Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID
1664792	40	1664792CD1	84	1664792CB1
2079396	41	2079396CD1	85	2079396CB1
5390115	42	5390115CD1	86	5390115CB1
1403326	43	1403326CD1	87	1403326CB1
7690129	44	7690129CD1	88	7690129CB1

Table 2

Polypeptide SEQ ID NO:	Incyte Polypeptide ID	GenBank ID NO:	Probability Score	GenBank Homolog
1	2101688CD1	g971461	1.50E-141	UDP-GalNAc:polypeptide N-acetyl-galactosaminyl transferase [Homo sapiens] (White,T. et al. J. Biol. Chem. (1995) 270(41):24156-65)
2	5452330CD1	g1139548	0	Seizure-related gene product 6 type 2 precursor [Mus musculus](Shimizu-Nishikawa, K. et al. (1995) Biochem. Biophys. Res. Commun. 216:382-389)
3	4362432CD1	g4467410	5.20E-40	Lysozyme [Gallus gallus] (Nakano, T. & Graf, T.(1992) Oncogene 7:527-534)
4	5308104CD1	g3878261	2.10E-92	Similarity to S. Pombe BEM1/BUD5 [Caenorhabditis elegans]
16	6436791CD1	g13274582	5.00E-39	Thymus atrophy-related protein [Mus musculus]
17	1820972CD1	g33702	2.20E-106	Immunoglobulin lambda light chain [Homo sapiens]
18	3286805CD1	g431420	1.50E-283	Macrophage specific protein MPS1 [Mus musculus] (Spilsbury, K. et al. (1995) Blood 85:1620-1629)
19	3506590CD1	g577056	1.00E-211	C gamma 3 [Homo sapiens]
29	1899010CD1	g13384378	8.00E-43	Putative phosphate translocator [Oryza sativa]
36	1957517CD1	g1572802	2.90E-65	Enterococcus faecalis TRAB [Caenorhabditis elegans]
37	866038CD1	g849238	1.90E-30	Similar to polyposis locus protein 1 [Caenorhabditis elegans]
38	3869704CD1	g33718	5.20E-114	Immunoglobulin lambda light chain [Homo sapiens]
43	1403326CD1	g3983152	8.10E-56	Schlafen3 Lymphoid growth regulatory protein [Mus musculus] (Schwarz, D.A. et al. (1998) Immunity 9:657-668)
44	7690129CD1	g6715117	3.10E-219	MTR1 [Homo sapiens] Melastatin/TRP related protein found in Beckwith-Wiedemann syndrome chromosomal region 11p15.5 (Prawitt, D. et al. (2000) Hum. Mol. Genet. 9:203-216)

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
1	2101688CD1	552	S200 S241 S313 S387 S399 S433 S45 S507 S84 S89 T130 T196 T237 T27 T35 T355 T41 T467 T5 Y408 Y74		Glycosyl transferase: S114-F292	HMMER_PFAM
					Glycosyl transferase: I147-D157 (P<0.021)	BLIMPS_PFAM
					PD003162:NACETYLGALACTOSAMINYLTRANSFERASE TRANSFERASE POLYPEPTIDE ACETYLGALACTOSAMINYLTRANSFERASE UDPGALNAC:POLYPEPTIDE GLYCOSYLTRANSFERASE PROTEINUDP PROTEIN UDP N: Q256-P414	BLAST_PRODOM
					ACETYLGALACTOSAMINYLTRANSFERASE; POLYPEPTIDE; DM03891|I37405|21-571: V26-W547	BLAST_DOMO
					Signal_peptide: M1-R29	HMMER
					Transmembrane domain: L4-W25	HMMER
					Signal_cleavage: M1-R29	SPSCAN
2	5452330CD1	994	S218 S249 S257 S263 S291 S378 S463 S501 S674 S724 S770 S780 S786 S820 S824 S842 S877 S919 S974 T38 T425 T553 T63 T647 T655 T709 T757 T812	N247 N289 N313 N399 N422 N436 N440 N541 N583 N707	Signal peptide: M1-G19	HMMER
					Signal peptide: M1-G19	SPSCAN
					Transmembrane domain: I930-Y947	HMMER
					Sushi domains (SCR repeats): C357-C412, C532-C589, C710-C765, C771-C830, C838-C895	HMMER_PFAM
					CUB domains: C416-Y524, C593-F701	HMMER_PFAM
					SEIZURERELATED GENE PRODUCT PRECURSOR SIGNAL TYPE PD024762:H18-A415 PD028803:V911-G984	BLAST_PRODOM
					SUSHI REPEAT DM04887|P33730|1-610: T735-D901, F381-P450, T548-I631 DM04887|P16581|1-609: T732-Y904, L354-P450, E525-P610 DM04887|P27113|1-551: S722-R896, L354-P450	BLAST_DOMO

Table 3 (cont.)

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
3	4362432CD1	212	S181 S190 S211 S26 T153 T16 T188 T45		Signal_cleavage: M1-G19	SPSCAN
					Signal_peptide: M1-G19	HMMER
					Transglycosylase SLT domain SLT: T82-A202	HMMER_PFAM
					Pterin 4 alpha carbinolamine dehydratase PF01329: G124-K130	BLIMPS_PFAM
					LYSOZYME G SIGNATURE PR00749: G174-D195, D191-S211, C39-M59, N60-Q81, I84-I102, S103-F123, G124-K142, K157-K173	BLIMPS_PRINTS
					LYSOZYME G 4BETA-N-ACETYLMURAMIDASE GOOSETYPE HYDROLASE PD016787: G38-F212	BLAST_PRODOM
					LYSOZYME G DM07376|P00718|1-184: C39-F212 DM07376|P27042|27-210: G38-F212	BLAST_DOMO
4	5308104CD1	308	S154 S158 S201 S5 S79 S93 T225 T253 T55 T71 Y163		Signal_cleavage: M1-G61	SPSCAN
					Dienelactone hydrolase family DL: P235-H262	HMMER_PFAM
					Tonb_Dependent_Receptor protein signature M1-S5	MOTIFS
					PROTEIN INTERGENIC REGION TRANSMEMBRANE OF TRAXFINO PLASMID SECTION BEM46 KRE1HXT14 PD009919: T113-S216	BLAST_PRODOM
					HYPOTHETICAL 34.9 KD PROTEIN HYPOTHETICAL PROTEIN PD126088: F234-S302	BLAST_PRODOM
					K04G2.2 PROTEIN PD126091: N2-E40	BLAST_PRODOM
5	3092736CD1	328	S116 S121 S148 S155 S159 S221 S278 S317 S52 T57		Signal_peptide: M1-A19	HMMER
					Signal_cleavage: M1-G22	SPSCAN
6	3580257CD1	69	T58		Signal_cleavage: M1-A21	SPSCAN
7	3634758CD1	158	T34 T55		Signal_cleavage: M1-G17	SPSCAN
8	4027923CD1	463	S113 S175 S360 S45 S86 T132 T157		Signal_peptide: M1-R37	HMMER
9	4348533CD1	648	S179 S244 S265 S303 S327 S329 S337 S389 S551 S571 S586 S620 S639 T276 T425 T470 T49 T496 T599 T606	N161 N310 N313	Signal_cleavage: M1-N68	SPSCAN
					Leucine_Zipper: L178-L199	MOTIFS

Table 3 (cont.)

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
10	4521857CD1	130	S10 T75		Signal_cleavage: M1-A38	SPSCAN
					Transmembrane domain: G20-Y40	HMMER
11	4722253CD1	279	S171 S230 S73 S77 T107 T243 T268	N191 N266 N71	Signal_cleavage: M1-A62	SPSCAN
12	4878134CD1	458	S15 S229 S279 S321 S340 S381 S439 T127 T93	N198 N259 N319	Transmembrane domain: L22-L41	HMMER
					Rgd: R118-D120	MOTIFS
13	5050133CD1	173	S130 S50		Signal_cleavage: M1-A31	SPSCAN
14	5630124CD1	335	S142 S191 S219 S295 S302 S324 S67 S74 T104 T190 T225 T243 T252 T275 T292 Y332		Signal_peptide: M1-A39	HMMER
					Signal_cleavage: M1-G36	SPSCAN
15	5677286CD1	71	T42		Signal_peptide: M1-A34	HMMER
					Signal_cleavage: M1-A66	SPSCAN
16	6436791CD1	148	S143 S16 T18	N31	Transmembrane domain: L109-F126	HMMER
17	1820972CD1	231	S140 S206 S219 S74		Signal_peptide: M1-S20	HMMER
					Signal_cleavage: M1-G16	SPSCAN
					do IMMUNOGLOBULIN; IG; HISTOCOMPATIBILITY; MAJOR DM02680|A39949|1-118: V115-C230	BLAST_DOMO
					MHC FRAMEWORK DOMAIN DM00397|S24319|1-128: M1-P128	BLAST_DOMO
					B-cell mu chain associated 8HS20 protein precursor PD174509: L23-V108	BLAST_PRODOM
					Immunoglobulins and MHC protein signature BL00290: T150-S172, Y210-P227	BLIMPS_BLOCKS
					Immunoglobulins and MHC protein signature ig_mhc.prf: K190-S231	PROFILESCAN
					Immunoglobulin domain ig: G34-V108, A146-V214	HMMER_PFAM
					Ig_Mhc: Y210-H216	MOTIFS
18	3286805CD1	716	S179 S231 S268 S331 S484 S553 S92 T147 T158 T207 T440 T447 T613 T679 T707 S19 T72 Y67 Y78	N185 N255 N269 N272 N375	Signal_peptide: M1-P22	HMMER
					Transmembrane domain: S653-I676	HMMER
					Signal_cleavage: M1-A17	SPSCAN

Table 3 (cont.)

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
19	3506590CD1	519	S104 S144 S339 S36 S396 S426 S75 S82 T234 T371 T509 Y113 Y368	N369	Signal_cleavage: M1-C19	SPSCAN
					Signal_peptide: M1-C19	HMMER
					MHC HINGE DOMAIN DM01060|P01862|1-329: S142-K275, R285-G518	BLAST_DOMO
					IG GAMMA3 CHAIN C REGION HEAVY DISEASE PROTEIN HDC IMMUNOGLOBULIN GLYCOPROTEIN PD028815: E241-G309	BLAST_PRODOM
					Immunoglobulins and MHC protein signature BL00290: S436-Q458, F495-S512	BLIMPS_BLOCKS
					Immunoglobulins and MHC protein signature ig_mhc.prf: T371-V420, D473-K519	PROFILESCAN
					Immunoglobulin domain ig: G34-R117, G162-V227, S326-V395, K432-V499	HMMER_PFAM
					Ig_Mhc: Y223-H229, F495-H501	MOTIFS
20	003600CD1	172	T73 T71 T90 S128		Signal peptide:M6-L26	HMMER
					Signal cleavage:M1-A28	SPSCAN
					Transmembrane domain:L12-N30	HMMER
					Leucine zipper motif:L12-L33	MOTIFS
21	1251534CD1	314			Signal peptide:M43-M67	HMMER
					Transmembrane domain:A250-I267	HMMER
22	1402211CD1	542	S430 S131 S137 S186 T273 S371 S395 T417 T426 S454 T34 S44 T114 S319 T509	N2 N359 N408 N409 N424 N529	Signal peptide:M345-H366	HMMER
23	1623474CD1	715	T66 S121 T216 T334 S376 S380 S386 T436 T475 T524 S543 S585 S586 S647 T659 S704 T709 S5 T108 T222 T279 S372 S390 S395 S406 S429 S445 S455 S503 S590 S639	N238 N335 N61 N239 N461 N465 N535	Rgd motif:R377-D379	MOTIFS
					Signal peptide:M187-V211	HMMER
					Transmembrane domain:I49-F67	HMMER

Table 3 (cont.)

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
24	1706443CD1	469	Y228 T70 S102 S158 T283 T337 S364 T37 S168 S179 T182 S292 S316 S359 T436 S462 S466		Rgd motif:R119-D121	MOTIFS
					Signal peptide:M1-G24	HMMER
					Signal cleavage:M1-G24	SPSCAN
25	1748627CD1	274	T9 T90 T237 S241 S248 S62 S100 S136 S191 T35	N254 N270	Signal cleavage:M1-A59	SPSCAN
26	1818332CD1	154	S120 S136 T41 S56 T76 S98 S138		Signal cleavage:M1-A26	SPSCAN
27	1822832CD1	102	T13 T19	N16	Signal peptide:M34-P57	HMMER
					Rgd motif:R21-D23	MOTIFS
28	1832219CD1	113			Signal cleavage:M1-G29	SPSCAN
29	1899010CD1	313	S127 S145 S300	N43 N92 N97 N98 N238	Signal peptide:M194-G211	HMMER
					Transmembrane domain:H11-I35, F151-V171, W219-V237	HMMER
30	2008768CD1	195	S35 S49 T64 S78 S117		Signal peptide:M121-A139	HMMER
					Transmembrane domain:L95-R116, N122-L145	HMMER
31	2070984CD1	350	T77	N294	Signal cleavage:M1-A66	SPSCAN
					Transmembrane domain:Y40-G61, M84-C102, V173-V191	HMMER
32	2193240CD1	360	Y327 S220 S221 S7 S38 T135 S318	N159 N207 N218 N142	Signal peptide:M101-S121	HMMER
33	2235177CD1	559	S301 S412 S520 T11 T27 S29 S42 T76 T156 S165 S252 T277 T303 T336 T462 T120 T121 S292 S322 S397 T407 T418	N70 N171 N357 N325 N417	Signal peptide:M191-A209	HMMER
34	2416227CD1	198	S136 S167 S137	N38 N68 N75 N92	Signal peptide:M1-S18	HMMER
					Signal cleavage:M1-S18	SPSCAN
					Transmembrane domain:F113-L133	HMMER
35	2461076CD1	73	T40 S25 T41		Signal peptide:M1-G21	HMMER
					Signal cleavage:M1-V19	SPSCAN

Table 3 (cont.)

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
36	1957517CD1	376	S87 T94 T196 S257 S326 S38 S224 S280	N36 N307		MOTIFS
37	866038CD1	216	T11 T15 S59 S114 S142 T146 S167 S172 S107 S157 T200 T209 S210 Y68		Leucine zipper motif:L129-L150	MOTIFS
					Signal cleavage:M1-G45	SPSCAN
38	3869704CD1	233	S112 S142 S208 S221 S74 T15 T36		Signal peptide: M1-A19	HMMER
					Signal cleavage: M1-A19	SPSCAN
					Immunoglobulins and major histocompatibility domains: Y212-H218	MOTIFS
					Immunoglobulins and major histocompatibility domains ig_mhc.prf: N191-S233	PROFILESCAN
					Immunoglobulins and major histocompatibility domains BL00290: T152-S174, Y212-P229	BLIMPS_BLOCKS
					Immunoglobulin domain ig: G34-S108, A148-V216	HMMER_PFAM
					IMMUNOGLOBULIN; MAJOR HISTOCOMPATIBILITY DM02680|A39949|1-118: V117-C232	BLAST_DOMO
					Immunoglobulin framework domain DM00397|S30526S|1-119: S20-F139	BLAST_DOMO
					IMMUNOGLOBULIN DM00001|S29258|119-206: T137-K225	BLAST_DOMO
39	1415179CD1	163	T104 T86		Signal cleavage: M1-S35	SPSCAN
					Mitochondrial Carrier: P134-M142	MOTIFS
					ZP receptor-type domain BL00682: C50-L56	BLIMPS_BLOCKS
40	1664792	235	S33 T70 T93 T94 T121 T224		Signal peptide M1-D18	HMMER
41	2079396CD1	94	S21 S45		Signal cleavage: M1-S42	SPSCAN
					GTP-binding elongation factors signature efactor_gtp.prf: M1-S52	PROFILESCAN
					Peroxidases signatures peroxidase_2.prf: I37-W90	PROFILESCAN
42	5390115CD1	85	S3 S8 T16 T63 T81		Signal cleavage: M1-S47	SPSCAN
					Transmembrane domain: Y24-I44	HMMER

Table 3 (cont.)

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
43	1403326CD1	901	S120 S13 S139 S219 S269 S383 S521 S531 S588 S603 S641 S708 S80 S805 S853 S858 T154 T230 T25 T296 T344 T352 T354 T493 T505 T650 T688 T776 T795 T815 Y279 Y311 Y681 Y804 Y824		P-loop Atp_Gtp_A: G599-T606	MOTIFS
44	7690129CD1	1040	S191 S254 S367 S539 S579 S679 S969 S971 S978 T112 T140 T182 T503 T535 T544 T729 T93	N116 N54 N818	Leucine_Zipper: L695-L716	MOTIFS
					Rgd: R40-D42 R241-D243	MOTIFS
					Transmembrane domain: V606-F623, M753-A773, W844-V862	HMMER
					PROTEIN CHROMOSOME TRANSMEMBRANE MELASTATIN C05C12.3 T01H8.5 I F54D1.5 IV PD151509: V730-A1018 PD018035: K8-W246 PD039592: Q382-E546	BLAST_PRODOM

Table 4

Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID	Sequence Length	Selected Fragment(s)	Sequence Fragments	5' Position	3' Position
45	2101688CB1	2508	71-123	6735891H1 (LIVRTUT13)	883	1375
				7620180J1 (KIDNTUE01)	1757	2290
				6874586H1 (EPIMUNN04)	1988	2508
				7704542H1 (UTRETUE01)	192	634
				3593046H1 (293TF5T01)	1	304
				6018547H1 (HNT2UNN03)	1327	2033
				6124211H1 (BRAHNON05)	1053	1632
				7700489J1 (KIDPTDE01)	409	938
46	5452330CB1	4034	1493-1673, 1-1081, 2638-2908, 3129-3535	3470968F6 (BRAIDIT01)	1862	2432
				6982855F8 (BRAIFER05)	1	427
				4775091H1 (BRAQNOT01)	3798	4034
				5404047T6 (BRAHNOT01)	3471	4026
				7293087F8 (BRAIFER06)	526	1205
				7583209H1 (BRAIFEC01)	301	861
				6207435H1 (PITUNON01)	3033	3733
				5404047F6 (BRAHNOT01)	2931	3445
				7115489H1 (BRAENOK01)	2298	2745
				6990568H1 (BRAIFER05)	480	1092
				7293087R8 (BRAIFER06)	1078	1790
				7579594H1 (BRAIFEC01)	2446	2962
				7291338F8 (BRAIFER06)	1717	2352
47	4362432CB1	845	1-44, 685-845	4362432F6 (SKIRNOT01)	1	664
				4362432T9 (SKIRNOT01)	228	845
48	5308104CB1	2300	1-807, 2192-2300	71013085V1	1689	2273
				6809635J1 (SKIRNOR01)	1	532
				8044501J1 (OVARTUE01)	214	765
				1550768R6 (PROSNOT06)	1989	2283
				6804176H1 (COLENOR03)	1230	1814
				71014150V1	656	1210
				6880707H1 (BRAHTDR03)	1100	1808
				503680H1 (TMLR3DT02)	2119	2300
49	3092736CB1	1587	1-180	SCGA02766V1	367	1073
				SCGA07870V1	685	1131
				1611754F6 (COLNTUT06)	1153	1587

Table 4 (Cont.)

Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID	Sequence Length	Selected Fragment(s)	Sequence Fragments	5' Position	3' Position
49				2823991F6 (ADRETUT06)	1031	1532
				SCGA12762V1	1	524
50	3580257CB1	669	1-24	3580257F6 (293TF3T01)	133	669
				5107219H1 (PROSTUS19)	1	240
				g1496797	1	495
51	3634758CB1	1463	1-51	4719037H1 (BRAIHCT02)	1177	1432
				SXAF05002V1	1	521
				SXAF05483V1	379	868
				3243342H1 (BRAINOT19)	1231	1463
				2729881H1 (OVARTUT04)	1095	1333
				SXAF05152V1	604	1131
52	4027923CB1	1686	1-204, 1666-1686	2532289H1 (GBLANOT02)	963	1179
				1281432F6 (COLNNOT16)	620	1173
				2561353H1 (ADRETUT01)	1	276
				664136H1 (SCORNOT01)	1428	1686
				3585158H1 (293TF4T01)	325	639
				1281432T6 (COLNNOT16)	1051	1684
				6772967J1 (BRAUNOR01)	75	607
53	4348533CB1	2497	1556-1848, 1-150, 2371-2497, 762-909	6933091H1 (SINTTMR02)	1346	1901
				g1617775	1	405
				2890155F6 (LUNGFET04)	1	483
				6781002J1 (OVARDIR01)	153	903
				2622331H1 (KERANOT02)	2139	2497
				2507578T6 (CONUTUT01)	1684	2359
				1728133F6 (PROSNOT14)	546	1160
				6945931H1 (FTUBTUR01)	1132	1795
54	4521857CB1	1783	1-733, 805-890	3003172H1 (TLYMNOT06)	900	1194
				4521857F6 (HNT2TXT01)	1	537
				825638T1 (PROSNOT06)	1100	1762
				857689R1 (NGANNOT01)	1194	1777
				3644845F6 (LUNGNOT34)	490	903
				362417R6 (PROSNOT01)	1227	1783
55	4722253CB1	1461	1-499	4722253H1 (COLCTUT02)	933	1204
				7018504H1 (KIDNNOC01)	1	668
				2455753F6 (ENDANOT01)	577	1109

Table 4 (Cont.)

Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID	Sequence Length	Selected Fragment(s)	Sequence Fragments	5' Position	3' Position
55				g3053012	975	1461
				3028265F6 (HEARFET02)	1292	1461
56	4878134CB1	2116	1-1071	3396235H1 (BRAIDIT01)	1865	2116
				4501019F6 (BRAVTXT02)	544	1119
				SBQA01857D1	467	1044
				3521653T6 (LUNGNON03)	1181	1713
				5021921H1 (OVARNON03)	1745	2041
				3766951H1 (BRSTNOT24)	1553	1851
				4501019T6 (BRAVTXT02)	1050	1704
				70874715V1	1	544
57	5050133CB1	702	1-28, 651-702	g1802638	321	702
				6871022H1 (BRAGNON02)	1	630
				3729290F6 (SMCCNON03)	199	686
58	5630124CB1	2613	1-975	6821390J1 (SINTNOR01)	520	1293
				6431012H1 (LUNGNON07)	1175	1881
				6855495H1 (BRAIFEN08)	1	643
				3878611T6 (SPLNNOT11)	2075	2590
				1358001T6 (LUNGNOT09)	1940	2586
				481430R7 (LIVRBCT01)	854	1381
				2252822R6 (OVARTUT01)	2311	2613
				481430T7 (LIVRBCT01)	1371	2013
59	5677286CB1	1778	1736-1778, 1-143, 672-767	70613827V1	1195	1777
				7053934H2 (BRACNOK02)	602	1280
				6340571H1 (BRANDIN01)	669	1324
				3620887T6 (BRSTNOT24)	1	642
				1810961F6 (PROSTUT12)	1421	1778
60	6436791CB1	1234	1-192	1212854T6 (BRSTTUT01)	556	1221
				3510032F6 (CONCNOT01)	1	590
				3943483F6 (SCORNOT04)	764	1234
61	1820972CB1	863	1-228, 843-863	60144357B1	227	833
				70636975V1	253	863
				1820972H1 (GBLATUT01)	1	267
62	3286805CB1	2521	1-155, 1165-2294	5030319F7 (COLCDIT01)	1300	1973
				7168560H1 (MCLRNOC01)	575	1020
				6959075H1 (SKINDIA01)	1	674

Table 4 (Cont.)

Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID	Sequence Length	Selected Fragment(s)	Sequence Fragments	5' Position	3' Position
62				3286805F6 (HEAONOT05)	1706	2266
				6466032H1 (PLACFEB01)	764	1319
				71054005V1	1949	2521
63	3506590CB1	1765	1-798	71409670V1	1051	1765
				7710638H1 (TESTTUE02)	524	1080
				70515763V1	1	556
				7733848H2 (COLDDIE01)	543	1266
64	003600CB1	1264	1-699	2718319H1 (THYRNOT09)	195	442
				2397316T6 (THP1AZT01)	714	1264
				003600R6 (HMC1NOT01)	509	1042
				008108H1 (HMC1NOT01)	269	498
				4592276H1 (MASTTXT01)	398	639
				2911566H1 (KIDNTUT15)	1	265
65	1251534CB1	3415	1-122, 2125-2328, 2982-3415, 834-1643	531771T6 (BRAINOT03)	2338	2967
				4697183F6 (BRALNOT01)	235	863
				487800H1 (HNT2AGT01)	1633	1906
				4740283H1 (THYMNOR02)	1210	1470
				1717327T6 (UCMCNOT02)	2865	3415
				7761153H1 (THYMNOE02)	1	708
				6855869H1 (BRAIFEN08)	1969	2594
				6442673H1 (BRAENOT02)	1814	2453
				3291485F6 (BONRFET01)	798	1410
				5404331H1 (BRAHNOT01)	1669	1969
				1251534H1 (LUNGFET03)	1448	1681
66	1402211CB1	2289	1707-2289	1650008F6 (PROSTUT09)	1296	1912
				429727T6 (BLADNOT01)	1626	2289
				2862734H1 (SININOT03)	1082	1342
				2129033R6 (KIDNNOT05)	1	512
				2499655F7 (ADRETUT05)	564	1084
				5397049H1 (LIVRTUT13)	1031	1293
				826301R1 (PROSNOT06)	1332	1928
				429727R6 (BLADNOT01)	335	862
67	1623474CB1	4480	2411-3066, 1-22, 109-587, 3638-3733	2158031F6 (BRAINOT09)	3973	4480

Table 4 (Cont.)

Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID	Sequence Length	Selected Fragment(s)	Sequence Fragments	5' Position	3' Position
67				3332425T6 (BRAIFET01)	2946	3643
				6914395J1 (PITUDIR01)	2128	2758
				6918276H1 (PLACFER06)	218	977
				6128115H1 (BRAHNON05)	2348	3030
				70758295V1	1025	1574
				3394074H1 (LUNGNOT28)	1	287
				1864803T6 (PROSNOT19)	3133	3718
				1975441T6 (PANCTUT02)	3747	4451
				1975441F6 (PANCTUT02)	3664	4246
				70760953V1	426	1008
				70761097V1	821	1416
				70757930V1	1651	2163
				60205344U1	1374	2060
68	1706443CB1	1568	1-43	6630210U1	438	905
				1858593T6 (PROSNOT18)	1044	1521
				1706443T6 (DUODNOT02)	906	1513
				1858593F6 (PROSNOT18)	686	989
				7062677H1 (PENITMN02)	1	468
				1390249H1 (EOSINOT01)	1342	1568
69	1748627CB1	1887	1-649	5407812F8 (BRAMNOT01)	429	985
				71427502V1	1299	1887
				6886749J1 (BRAHTDR03)	632	1022
				3627391F6 (COLNNOT38)	1	531
				71430251V1	1047	1612
				1979283R6 (LUNGTUT03)	997	1556
70	1818332CB1	569	1-35	1255779F2 (MENITUT03)	39	569
				1336021T1 (COLNNOT13)	1	541
71	1822832CB1	2338	529-565, 1332-1369, 1-124, 1488-2338	1289709F6 (BRAINOT11)	1977	2338
				1822832X352U1 (GBLATUT01)	25	655
				g1975312	1	277
				SAOA01720F1	710	1312
				SAOA01416F1	1318	1894
				1452843F6 (PENITUT01)	520	1212
				SAOA01295F1	1707	2338
				SAOA00837F1	1223	1841

Table 4 (Cont.)

Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID	Sequence Length	Selected Fragment(s)	Sequence Fragments	5' Position	3' Position
72	1832219CB1	481	1-21	SXAF02203V1	1	479
				1832219R6 (BRAINON01)	44	481
73	1899010CB1	1255	1-62	1899010F6 (BLADTUT06)	341	824
				2174773F6 (ENDCNOT03)	863	1255
				1909527T6 (CONNTUT01)	569	1233
				1425473H1 (BEPINON01)	1	264
				1909527F6 (CONNTUT01)	34	643
74	2008768CB1	875	1-411	2008768T6 (TESTNOT03)	159	858
				6025379H1 (TESTNOT11)	1	261
				563323R6 (NEUTLPT01)	472	875
75	2070984CB1	2188	1-72, 1579-1656	1273987F1 (TESTTUT02)	1769	2188
				7162982H1 (PLACNOR01)	321	914
				SBIA08036D1	1307	1793
				SBIA01466D1	891	1516
				6907510J1 (PITUDIR01)	112	851
				3320716H1 (PROSBPT03)	1	280
76	2193240CB1	1561	1-624	1624251F6 (BRAITUT13)	1	472
				2429918R6 (MENTUNON2)	734	1174
				2429918T6 (MENTUNON2)	1003	1561
				900981R1 (BRSTTUT03)	365	926
77	2235177CB1	1777	1-32	71113502V1	1140	1777
				6993448H1 (BRAQTDR02)	1	725
				71264559V1	837	1353
				71113614V1	679	1308
78	2416227CB1	1841	1-482	6821668J1 (SINTNOR01)	194	980
			518-1018	2416227T6 (HNT3AZT01)	1210	1789
				2416227F6 (HNT3AZT01)	976	1434
				854765H1 (NGANNOT01)	774	1005
				7017947H1 (KIDNNOC01)	1	604
				7154302H1 (HEARNONO3)	1446	1841
79	2461076CB1	1616	835-861, 565-783, 1-219	7039965H1 (UTRSTMR02)	110	673
				219625R6 (STOMNOT01)	948	1492
				6935657H1 (SINTTMR02)	855	1446
				219625T6 (STOMNOT01)	958	1616
				2461076F6 (THYRNOT08)	484	938
				6073858H1 (UTREDIT09)	1	273

Table 4 (Cont.)

Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID	Sequence Length	Selected Fragment(s)	Sequence Fragments	5' Position	3' Position
80	1957517CB1	1434	1-111	7161288H1 (PLACNOR01)	448	1014
				1233279F1 (LUNGFET03)	1	537
				6573238H1 (COLHTUS02)	655	1348
				1575944H1 (LNODNOT03)	1214	1434
81	866038CB1	2085	51-124	6913288J1 (PITUDIR01)	272	823
				5964475H1 (BRATNOT05)	1744	2085
				7132506H1 (BRAHTDK01)	816	1499
				6054811H1 (BRAENOT04)	784	1447
				755563R1 (BRAITUT02)	215	753
				5483139H1 (FIBPFEN06)	1	285
				6438276H1 (BRAENOT02)	1420	2083
82	3869704CB1	904	1-36	705156H1 (SYNORAT04)	1	232
				71052048V1	331	904
				3901129R8 (LUNGNON03)	48	697
				3904253R9 (LUNGNON03)	217	837
83	1415179CB1	1496	1-248, 606-836	2042611R6 (HIPONON02)	729	1206
				660950R6 (BRAINOT03)	557	1155
				4713560H1 (BRAIHCT01)	1	252
				658639F1 (BRAINOT03)	839	1496
				2708523H1 (PONSAZT01)	430	732
				2967826F6 (SCORNOT04)	58	686
84	1664792CB1	2837	1-1559	70858742V1	407	959
				4797546H1 (LIVRTUT09)	1248	1524
				2699003T6 (OVARTUT10)	2154	2825
				71224728V1	867	1504
				2542259F6 (BONRTUT01)	1859	2386
				7460645H1 (LIVRTUE01)	441	1062
				2260182R6 (UTRSNOT02)	2545	2837
				1664792T6 (BRSTNOT09)	1698	2381
				6988160H1 (BRAIFER05)	1	444
				1664792F6 (BRSTNOT09)	1257	1834
				4179737H1 (SINITUT03)	1589	1863
85	2079396CB1	1123	1-45, 993-1123	6820736H1 (SINTNOR01)	566	1063
				g1401473	1	507
				874769R1 (LUNGAST01)	786	1123
				6819702J1 (OVARDIR01)	54	794
85				6335288H1 (BRANDIN01)	17	509

Table 4 (Cont.)

Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID	Sequence Length	Selected Fragment(s)	Sequence Fragments	5' Position	3' Position
86	5390115CB1	1549	1-270	1258145F1 (MENITUT03)	589	1247
				1466677F1 (PANCTUT02)	1086	1549
				4250616F6 (BRADDIR01)	1	633
				1310308F1 (COLNFET02)	758	1337
87	1403326CB1	4820	1-3502	1306452F6 (PLACNOT02)	1069	1588
				70607520V1	291	789
				4322557H1 (TLYMUNT01)	2789	3045
				3080429F6 (BRAIUNT01)	2387	3023
				70476331V1	1812	2459
				70604815V1	1	395
				2801448F6 (PENCNOT01)	3054	3590
				4729710H1 (GBLADIT01)	1585	1844
				6245574H1 (TESTNOT17)	3553	4124
				70815905V1	417	793
				5718724H1 (PANCNOT16)	1188	1815
				6863174H1 (BRAGNON02)	4389	4820
				6937903H1 (FTUBTUR01)	3745	4307
				6489460H1 (MIXDUNB01)	2100	2703
				4884473H2 (LUNLTMT01)	2991	3242
				2820527T6 (BRSTNOT14)	4094	4505
				5642645R8 (UTRSTMR01)	603	1080
88	7690129CB1	3599	1878-1968, 1-934, 2349-3111	1251961F1 (LUNGFET03)	2550	3112
				1851125T6 (LUNGFET03)	2986	3599
				7757184J1 (SPLNTUE01)	706	1447
				6800356J1 (COLENOR03)	222	920
				6831480J1 (SINTNOR01)	1821	2515
				6883161H1 (BRAHTDR03)	580	1029
				5868845F8 (COLTDIT04)	1644	2425
				71137279V1	2150	2848
				2185757F6 (PROSNOT26)	1	495
				7612578J1 (KIDCTME01)	1041	1732

Table 5

Polynucleotide SEQ ID NO:	Incyte Project ID	Representative Library
45	2101688CB1	BRAITUT02
46	5452330CB1	BRAIDIT01
47	4362432CB1	SKIRNOT01
48	5308104CB1	BRAYDIN03
49	3092736CB1	BRAITUT08
50	3580257CB1	293TF3T01
51	3634758CB1	HUVENOB01
52	4027923CB1	COLNNOT16
53	4348533CB1	LIVRNON08
54	4521857CB1	SPLNNOT04
55	4722253CB1	TESTNOT03
56	4878134CB1	LUNGNON03
57	5050133CB1	FIBPFEN06
58	5630124CB1	LUNGNOT09
59	5677286CB1	PROSTUT12
60	6436791CB1	MEGBUNT01
61	1820972CB1	SPLNNOT04
62	3286805CB1	SKINDIA01
63	3506590CB1	COLDDIE01
64	003600CB1	HMC1NOT01
65	1251534CB1	THYMNOT05
66	1402211CB1	CARCTXT02
67	1623474CB1	HMC1NOT01
68	1706443CB1	DUODNOT02
69	1748627CB1	FIBPFEN06
70	1818332CB1	ISLTNOT01
71	1822832CB1	BRAINOT11
72	1832219CB1	TESTNOT03
73	1899010CB1	BLADTUT06
74	2008768CB1	TESTNOT03
75	2070984CB1	PLACNOT07
76	2193240CB1	BRAITUT13
77	2235177CB1	HNT2AGT01

Table 5 (Cont.)

Polynucleotide SEQ ID NO:	Incyte Project ID	Representative Library
78	2416227CB1	LUNGNOT09
79	2461076CB1	STOMNOT01
80	1957517CB1	OVARTUT01
81	866038CB1	BRAITUT03
82	3869704CB1	LUNGNOT03
83	1415179CB1	BRAINOT03
84	1664792CB1	BRSTTUT01
85	2079396CB1	CONUTUT01
86	5390115CB1	BRAITUT03
87	1403326CB1	BRSTNOT01
88	7690129CB1	PROSTUT12

Table 6

Library	Vector	Library Description
293TF3T01	pINCY	Library was constructed using RNA isolated from a serum-starved transformed embryonal cell line (293-EBNA) derived from kidney epithelial tissue. The cells were transformed with adenovirus 5 DNA.
BLADTUT06	pINCY	Library was constructed using RNA isolated from bladder tumor tissue removed from the posterior bladder wall of a 58-year-old Caucasian male during a radical cystectomy, radical prostatectomy, and gastrostomy. Pathology indicated grade 3 transitional cell carcinoma in the left lateral bladder wall. The remaining bladder showed marked cystitis with scattered microscopic foci of transitional cell carcinoma in situ. Patient history included angina, emphysema and tobacco use. Family history included acute myocardial infarction, atherosclerotic coronary artery disease, and type II diabetes.
BRAIDIT01	pINCY	Library was constructed using RNA isolated from diseased brain tissue. Patient history included multiple sclerosis, type II lesion.
BRAINOT03	PSPORT1	Library was constructed using RNA isolated from brain tissue removed from a 26-year-old Caucasian male during cranioplasty and excision of a cerebral meningeal lesion. Pathology for the associated tumor tissue indicated a grade 4 oligoastrocytoma in the right fronto-parietal part of the brain.
BRAINOT11	pINCY	Library was constructed using RNA isolated from brain tissue removed from the right temporal lobe of a 5-year-old Caucasian male during a hemispherectomy. Pathology indicated extensive polymicrogyria and mild to moderate gliosis (predominantly subpial and subcortical), consistent with chronic seizure disorder. Family history included a cervical neoplasm.
BRAITUT02	PSPORT1	Library was constructed using RNA isolated from brain tumor tissue removed from the frontal lobe of a 58-year-old Caucasian male during excision of a cerebral meningeal lesion. Pathology indicated a grade 2 metastatic hypernephroma. Patient history included a grade 2 renal cell carcinoma, insomnia, and chronic airway obstruction. Family history included a malignant neoplasm of the kidney.
BRAITUT03	PSPORT1	Library was constructed using RNA isolated from brain tumor tissue removed from the left frontal lobe of a 17-year-old Caucasian female during excision of a cerebral meningeal lesion. Pathology indicated a grade 4 fibrillary giant and small-cell astrocytoma. Family history included benign hypertension and cerebrovascular disease.
BRAITUT08	pINCY	Library was constructed using RNA isolated from brain tumor tissue removed from the left frontal lobe of a 47-year-old Caucasian male during excision of cerebral meningeal tissue. Pathology indicated grade 4 fibrillary astrocytoma with focal tumoral radionecrosis. Patient history included cerebrovascular disease, deficiency anemia, hyperlipidemia, epilepsy, and tobacco use. Family history included cerebrovascular disease and a malignant prostate neoplasm.
BRAITUT13	pINCY	Library was constructed using RNA isolated from brain tumor tissue removed from the left frontal lobe of a 68-year-old Caucasian male during excision of a cerebral meningeal lesion. Pathology indicated a meningioma in the left frontal lobe.

Table 6 (Cont.)

Library	Vector	Library Description
BRAYDIN03	pINCY	This normalized library was constructed from 6.7 million independent clones from a brain tissue library. Starting RNA was made from RNA isolated from diseased hypothalamus tissue removed from a 57-year-old Caucasian male who died from a cerebrovascular accident. Patient history included Huntington's disease and emphysema. The library was normalized in 2 rounds using conditions adapted from Soares et al., PNAS (1994) 91:9228 and Bonaldo et al., Genome Research (1996) 6:791, except that a significantly longer (48 -hours/round) reannealing hybridization was used. The library was linearized and recircularized to select for insert containing clones.
BRSTNOT01	PBLUESCRIPT	Library was constructed using RNA isolated from the breast tissue of a 56-year-old Caucasian female who died in a motor vehicle accident.
BRSTTUT01	PSPORT1	Library was constructed using RNA isolated from breast tumor tissue removed from a 55-year-old Caucasian female during a unilateral extended simple mastectomy. Pathology indicated invasive grade 4 mammary adenocarcinoma of mixed lobular and ductal type, extensively involving the left breast. The tumor was identified in the deep dermis near the lactiferous ducts with extracapsular extension. Seven mid and low and five high axillary lymph nodes were positive for tumor. Proliferative fibrocysytic changes were characterized by apocrine metaplasia, sclerosing adenosis, cyst formation, and ductal hyperplasia without atypia. Patient history included atrial tachycardia, blood in the stool, and a benign breast neoplasm. Family history included benign hypertension, atherosclerotic coronary artery disease, cerebrovascular disease, and depressive disorder.
CARCTXT02	PSPORT1	Library was constructed using RNA from chondrocytes that were isolated from pooled knee cartilage obtained during total knee joint replacement. The cartilage was removed from the underlying bone, chopped into smaller pieces, and stimulated with 5 ng/ml IL-1 for 18 hours.
COLDDIE01	PCDNA2.1	This 5 prime biased random primed library was constructed using RNA isolated from diseased descending colon tissue removed from a 28-year-old Caucasian male during a total intra-abdominal colectomy and temporary ileostomy. Pathology indicated chronic ulcerative colitis, moderate to severe, actively involving the distal 23 cm of colon. The entire 24 cm segment of rectosigmoid, rectum, and rectal tissue was involved with chronic ulcerative colitis, severely active. The patient presented with blood in the stool, diarrhea, and deficiency anemia. Patient history included shoulder dystonia (sprained rotator cuff), and tobacco abuse. The patient was treated with a transfusion. Patient medications included Asacol, Prednisone, and cortisone enemas. Family history included acute myocardial infarction, upper lobe lung cancer, colon cancer, and type I diabetes in the grandparent(s).
COLNNOT16	pINCY	Library was constructed using RNA isolated from sigmoid colon tissue removed from a 62-year-old Caucasian male during a sigmoidectomy and permanent colostomy.
CONUTUT01	pINCY	Library was constructed using RNA isolated from sigmoid mesentery tumor tissue obtained from a 61-year-old female during a total abdominal hysterectomy and bilateral salpingo-oophorectomy with regional lymph node excision. Pathology indicated a metastatic grade 4 malignant mixed mullerian tumor present in the sigmoid mesentery at two sites.

Table 6 (Cont.)

Library	Vector	Library Description
DUODNOT02	pINCY	Library was constructed using RNA isolated from duodenal tissue of a 8-year-old Caucasian female, who died from head trauma. Serology was positive for cytomegalovirus (CMV).
FIBPFEN06	pINCY	The normalized prostate stromal fibroblast tissue libraries were constructed from 1.56 million independent clones from a prostate fibroblast library. Starting RNA was made from fibroblasts of prostate stroma removed from a male fetus,who died after 26 weeks' gestation. The libraries were normalized in two roundsusing conditions adapted from Soares et al., PNAS (1994) 91:9228 and Bonaldo et al., Genome Research (1996) 6:791, except that a significantly longer (48-hours/round)reannealing hybridization was used. The library was then linearized andrecircularized to select for insert containing clones as follows: plasmid DNA wasprepped from approximately 1 million clones from the normalized prostate stromalfibroblast tissue libraries following soft agar transformation.
HMC1NOT01	PBLUESCRIPT	Library was constructed using RNA isolated from the HMC-1 human mast cell line derived from a 52-year-old female. Patient history included mast cell leukemia.
HNT2AGT01	PBLUESCRIPT	Library was constructed at Stratagene (STR937233), using RNA isolated from the hNT2 cell line derived from a human teratocarcinoma that exhibited properties characteristic of a committed neuronal precursor. Cells were treated with retinoic acid for 5 weeks and with mitotic inhibitors for two weeks and allowed to mature for an additional 4 weeks in conditioned medium.
HUVENOB01	PBLUESCRIPT	Library was constructed using RNA isolated from HUV-EC-C (ATCC CRL 1730) cells.
ISLTNOT01	pINCY	Library was constructed using RNA isolated from a pooled collection of pancreatic islet cells.

Table 6 (Cont.)

Library	Vector	Library Description
LIVRNON08	pINCY	This normalized library was constructed from 5.7 million independent clones from a pooled liver tissue library. Starting RNA was made from pooled liver tissue removed from a 4-year-old Hispanic male who died from anoxia and a 16 week female fetus who died after 16-weeks gestation from anencephaly. Serologies were positive for cytolomegalovirus in the 4-year-old. Patient history included asthma in the 4-year-old. Family history included taking daily prenatal vitamins and mitral valve prolapse in the mother of the fetus. The library was normalized in 2 rounds using conditions adapted from Soares et al., PNAS (1994) 91:9228 and Bonaldo et al., Genome Research 6 (1996):791, except that a significantly longer (48 hours/round) reannealing hybridization was used.
LUNGNON03	PSPORT1	This normalized library was constructed from 2.56 million independent clones from a lung tissue library. RNA was made from lung tissue removed from the left lobe a 58-year-old Caucasian male during a segmental lung resection. Pathology for the associated tumor tissue indicated a metastatic grade 3 (of 4) osteosarcoma. Patient history included soft tissue cancer, secondary cancer of the lung, prostate cancer, and an acute duodenal ulcer with hemorrhage. Patient also received radiation therapy to the retroperitoneum. Family history included prostate cancer, breast cancer, and acute leukemia. The normalization and hybridization conditions were adapted from Soares et al., PNAS (1994) 91:9228; Swaroop et al., NAR (1991) 19:1954; and Bonaldo et al., Genome Research (1996) 6:791.
LUNGNOT03	PSPORT1	Library was constructed using RNA isolated from lung tissue of a 79-year-old Caucasian male. Pathology for the associated tumor tissue indicated grade 4 carcinoma. Patient history included a benign prostate neoplasm and atherosclerosis.
LUNGNOT09	pINCY	Library was constructed using RNA isolated from the lung tissue of a 23-week-old Caucasian male fetus. The pregnancy was terminated following a diagnosis by ultrasound of infantile polycystic kidney disease.
MEGBUNT01	pINCY	Library was constructed using RNA isolated from an untreated MEG-01 megakaryoblast cell line, derived from bone marrow cells obtained from a 55-year-old male in megakaryoblastic crisis of chronic myelogenous leukemia.
OVARTUT01	PSPORT1	Library was constructed using RNA isolated from ovarian tumor tissue removed from a 43-year-old Caucasian female during removal of the fallopian tubes and ovaries. Pathology indicated grade 2 mucinous cystadenocarcinoma involving the entire left ovary. Patient history included mitral valve disorder, pneumonia, and viral hepatitis. Family history included atherosclerotic coronary artery disease, pancreatic cancer, stress reaction, cerebrovascular disease, breast cancer, and uterine cancer.
PLACNOT07	pINCY	Library was constructed using RNA isolated from placental tissue removed from a Caucasian fetus, who died after 16 weeks' gestation from fetal demise and hydrocephalus. Serology was positive for anti-CMV (cytomegalovirus).

Table 6 (Cont.)

Library	Vector	Library Description
PROSTUT12	pINCY	Library was constructed using RNA isolated from prostate tumor tissue removed from a 65-year-old Caucasian male during a radical prostatectomy. Pathology indicated an adenocarcinoma (Gleason grade 2+2). Adenofibromatous hyperplasia was also present. The patient presented with elevated prostate specific antigen (PSA).
SKINDIA01	PSPORT1	This amplified library was constructed using RNA isolated from diseased skin tissue removed from 1 female and 4 males during skin biopsies. Pathologies indicated tuberculoid and lepromatious leprosy.
SKIRNOT01	pINCY	Library was constructed using RNA isolated from skin tissue removed from the breast of a 26-year-old Caucasian female during bilateral reduction mammoplasty.
SPLNNOT04	pINCY	Library was constructed using RNA isolated from the spleen tissue of a 2-year-old Hispanic male, who died from cerebral anoxia. Past medical history and serologies were negative.
STOMNOT01	PBLUESCRIPT	Library was constructed using RNA isolated from the stomach tissue of a 55-year-old Caucasian male, who died from cardiopulmonary arrest.
TESTNOT03	PBLUESCRIPT	Library was constructed using RNA isolated from testicular tissue removed from a 37-year-old Caucasian male, who died from liver disease. Patient history included cirrhosis, jaundice, and liver failure.
THYMNOT05	pINCY	Library was constructed using RNA isolated from thymus tissue removed from a 3-year-old Hispanic male during a thymectomy and closure of a patent ductus arteriosus. The patient presented with severe pulmonary stenosis and cyanosis. Patient history included a cardiac catheterization and echocardiogram. Previous surgeries included Blalock-Taussig shunt and pulmonary valvotomy. The patient was not taking any medications. Family history included benign hypertension, osteoarthritis, depressive disorder, and extrinsic asthma in the grandparent(s).

Table 7

Program	Description	Reference	Parameter Threshold
ABI FACTURA	A program that removes vector sequences and masks ambiguous bases in nucleic acid sequences.	Applied Biosystems, Foster City, CA.	
ABI/PARACEL FDF	A Fast Data Finder useful in comparing and annotating amino acid or nucleic acid sequences.	Applied Biosystems, Foster City, CA; Paracel Inc., Pasadena, CA.	Mismatch <50%
ABI AutoAssembler	A program that assembles nucleic acid sequences.	Applied Biosystems, Foster City, CA.	
BLAST	A Basic Local Alignment Search Tool useful in sequence similarity search for amino acid and nucleic acid sequences. BLAST includes five functions: blastp, blastn, blastx, tblastn, and tblastx.	Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410; Altschul, S.F. et al. (1997) Nucleic Acids Res. 25:3389-3402.	*ESTs:* Probability value= 1.0E-8 or less *Full Length sequences*: Probability value= 1.0E-10 or less
FASTA	A Pearson and Lipman algorithm that searches for similarity between a query sequence and a group of sequences of the same type. FASTA comprises as least five functions: fasta, tfasta, fastx, tfastx, and ssearch.	Pearson, W.R. and D.J. Lipman (1988) Proc. Natl. Acad Sci. USA 85:2444-2448; Pearson, W.R. (1990) Methods Enzymol. 183:63-98; and Smith, T.F. and M.S. Waterman (1981) Adv. Appl. Math. 2:482-489.	*ESTs:* fasta E value=1.06E-6 *Assembled ESTs*: fasta Identity= 95% or greater and Match length=200 bases or greater; fastx E value=1.0E-8 or less *Full Length sequences:* fastx score=100 or greater
BLIMPS	A BLocks IMProved Searcher that matches a sequence against those in BLOCKS, PRINTS, DOMO, PRODOM, and PFAM databases to search for gene families, sequence homology, and structural fingerprint regions.	Henikoff, S. and J.G. Henikoff (1991) Nucleic Acids Res. 19:6565-6572; Henikoff, J.G. and S. Henikoff (1996) Methods Enzymol. 266:88-105; and Attwood, T.K. et al. (1997) J. Chem. Inf. Comput. Sci. 37:417-424.	Probability value= 1.0E-3 or less
HMMER	An algorithm for searching a query sequence against hidden Markov model (HMM)-based databases of protein family consensus sequences, such as PFAM.	Krogh, A. et al. (1994) J. Mol. Biol. 235:1501-1531; Sonnhammer, E.L.L. et al. (1988) Nucleic Acids Res. 26:320-322; Durbin, R. et al. (1998) Our World View, in a Nutshell, Cambridge Univ. Press, pp. 1-350.	*PFAM hits:* Probability value= 1.0E-3 or less *Signal peptide hits:* Score= 0 or greater

Table 7 (cont.)

Program	Description	Reference	Parameter Threshold
ProfileScan	An algorithm that searches for structural and sequence motifs in protein sequences that match sequence patterns defined in Prosite.	Gribskov, M. et al. (1988) CABIOS 4:61-66; Gribskov, M. et al. (1989) Methods Enzymol. 183:146-159; Bairoch, A. et al. (1997) Nucleic Acids Res. 25:217-221.	Normalized quality score≥GCG-specified "HIGH" value for that particular Prosite motif. Generally, score=1.4-2.1.
Phred	A base-calling algorithm that examines automated sequencer traces with high sensitivity and probability.	Ewing, B. et al. (1998) Genome Res. 8:175-185; Ewing, B. and P. Green (1998) Genome Res. 8:186-194.	
Phrap	A Phils Revised Assembly Program including SWAT and CrossMatch, programs based on efficient implementation of the Smith-Waterman algorithm, useful in searching sequence homology and assembling DNA sequences.	Smith, T.F. and M.S. Waterman (1981) Adv. Appl. Math. 2:482-489; Smith, T.F. and M.S. Waterman (1981) J. Mol. Biol. 147:195-197; and Green, P., University of Washington, Seattle, WA.	Score= 120 or greater; Match length= 56 or greater
Consed	A graphical tool for viewing and editing Phrap assemblies.	Gordon, D. et al. (1998) Genome Res. 8:195-202.	
SPScan	A weight matrix analysis program that scans protein sequences for the presence of secretory signal peptides.	Nielson, H. et al. (1997) Protein Engineering 10:1-6; Claverie, J.M. and S. Audic (1997) CABIOS 12:431-439.	Score=3.5 or greater
TMAP	A program that uses weight matrices to delineate transmembrane segments on protein sequences and determine orientation.	Persson, B. and P. Argos (1994) J. Mol. Biol. 237:182-192; Persson, B. and P. Argos (1996) Protein Sci. 5:363-371.	
TMHMMER	A program that uses a hidden Markov model (HMM) to delineate transmembrane segments on protein sequences and determine orientation.	Sonnhammer, E.L. et al. (1998) Proc. Sixth Intl. Conf. on Intelligent Systems for Mol. Biol., Glasgow et al., eds., The Am. Assoc. for Artificial Intelligence Press, Menlo Park, CA, pp. 175-182.	
Motifs	A program that searches amino acid sequences for patterns that matched those defined in Prosite.	Bairoch, A. et al. (1997) Nucleic Acids Res. 25:217-221; Wisconsin Package Program Manual, version 9, page M51-59, Genetics Computer Group, Madison, WI.	

What is claimed is:

1. An isolated polypeptide selected from the group consisting of:

a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-44,

b) a naturally occurring polypeptide comprising an amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-44,

c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-44, and

d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-44.

2. An isolated polypeptide of claim 1 selected from the group consisting of SEQ ID NO:1-44.

3. An isolated polynucleotide encoding a polypeptide of claim 1.

4. An isolated polynucleotide encoding a polypeptide of claim 2.

5. An isolated polynucleotide of claim 4 selected from the group consisting of SEQ ID NO:45-88.

6. A recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide of claim 3.

7. A cell transformed with a recombinant polynucleotide of claim 6.

8. A transgenic organism comprising a recombinant polynucleotide of claim 6.

9. A method for producing a polypeptide of claim 1, the method comprising:

a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell is transformed with a recombinant polynucleotide, and said recombinant polynucleotide comprises a promoter sequence operably linked to a polynucleotide encoding the polypeptide of claim 1, and

b) recovering the polypeptide so expressed.

10. An isolated antibody which specifically binds to a polypeptide of claim 1.

11. An isolated polynucleotide selected from the group consisting of:

a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:45-88,

b) a naturally occurring polynucleotide comprising a polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:45-88,

c) a polynucleotide complementary to a polynucleotide of a),

d) a polynucleotide complementary to a polynucleotide of b), and

e) an RNA equivalent of a)-d).

12. An isolated polynucleotide comprising at least 60 contiguous nucleotides of a polynucleotide of claim 11.

13. A method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 11, the method comprising:

a) hybridizing the sample with a probe comprising at least 20 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization complex is formed between said probe and said target polynucleotide or fragments thereof, and

b) detecting the presence or absence of said hybridization complex, and, optionally, if present, the amount thereof.

14. A method of claim 13, wherein the probe comprises at least 60 contiguous nucleotides.

15. A method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 11, the method comprising:

a) amplifying said target polynucleotide or fragment thereof using polymerase chain reaction amplification, and

b) detecting the presence or absence of said amplified target polynucleotide or fragment thereof, and, optionally, if present, the amount thereof.

16. A composition comprising a polypeptide of claim 1 and a pharmaceutically acceptable excipient.

17. A composition of claim 16, wherein the polypeptide has an amino acid sequence selected from the group consisting of SEQ ID NO:1-44.

18. A method for treating a disease or condition associated with decreased expression of functional SECP, comprising administering to a patient in need of such treatment the composition of claim 16.

19. A method for screening a compound for effectiveness as an agonist of a polypeptide of claim 1, the method comprising:

a) exposing a sample comprising a polypeptide of claim 1 to a compound, and

b) detecting agonist activity in the sample.

20. A composition comprising an agonist compound identified by a method of claim 19 and a pharmaceutically acceptable excipient.

21. A method for treating a disease or condition associated with decreased expression of functional SECP, comprising administering to a patient in need of such treatment a composition of claim 20.

22. A method for screening a compound for effectiveness as an antagonist of a polypeptide of claim 1, the method comprising:

a) exposing a sample comprising a polypeptide of claim 1 to a compound, and

b) detecting antagonist activity in the sample.

23. A composition comprising an antagonist compound identified by a method of claim 22 and a pharmaceutically acceptable excipient.

24. A method for treating a disease or condition associated with overexpression of functional SECP, comprising administering to a patient in need of such treatment a composition of claim 23.

25. A method of screening for a compound that specifically binds to the polypeptide of claim 1, said method comprising the steps of:

a) combining the polypeptide of claim 1 with at least one test compound under suitable conditions, and

b) detecting binding of the polypeptide of claim 1 to the test compound, thereby identifying a compound that specifically binds to the polypeptide of claim 1.

26. A method of screening for a compound that modulates the activity of the polypeptide of claim 1, said method comprising:

a) combining the polypeptide of claim 1 with at least one test compound under conditions permissive for the activity of the polypeptide of claim 1,

b) assessing the activity of the polypeptide of claim 1 in the presence of the test compound, and

c) comparing the activity of the polypeptide of claim 1 in the presence of the test compound with the activity of the polypeptide of claim 1 in the absence of the test compound, wherein a change in the activity of the polypeptide of claim 1 in the presence of the test compound is indicative of a compound that modulates the activity of the polypeptide of claim 1.

27. A method for screening a compound for effectiveness in altering expression of a target polynucleotide, wherein said target polynucleotide comprises a sequence of claim 5, the method comprising:

a) exposing a sample comprising the target polynucleotide to a compound, under conditions suitable for the expression of the target polynucleotide,

b) detecting altered expression of the target polynucleotide, and

c) comparing the expression of the target polynucleotide in the presence of varying amounts of the compound and in the absence of the compound.

28. A method for assessing toxicity of a test compound, said method comprising:

a) treating a biological sample containing nucleic acids with the test compound;

b) hybridizing the nucleic acids of the treated biological sample with a probe comprising at least 20 contiguous nucleotides of a polynucleotide of claim 11 under conditions whereby a specific hybridization complex is formed between said probe and a target polynucleotide in the biological sample, said target polynucleotide comprising a polynucleotide sequence of a polynucleotide of claim 11 or fragment thereof;

c) quantifying the amount of hybridization complex; and

d) comparing the amount of hybridization complex in the treated biological sample with the amount of hybridization complex in an untreated biological sample, wherein a difference in the

amount of hybridization complex in the treated biological sample is indicative of toxicity of the test compound.

29. A diagnostic test for a condition or disease associated with the expression of SECP in a biological sample comprising the steps of:

a) combining the biological sample with an antibody of claim 10, under conditions suitable for the antibody to bind the polypeptide and form an antibody:polypeptide complex; and

b) detecting the complex, wherein the presence of the complex correlates with the presence of the polypeptide in the biological sample.

30. The antibody of claim 10, wherein the antibody is:

a) a chimeric antibody,

b) a single chain antibody,

c) a Fab fragment,

d) a $F(ab')_2$ fragment, or

e) a humanized antibody.

31. A composition comprising an antibody of claim 10 and an acceptable excipient.

32. A method of diagnosing a condition or disease associated with the expression of SECP in a subject, comprising administering to said subject an effective amount of the composition of claim 31.

33. A composition of claim 31, wherein the antibody is labeled.

34. A method of diagnosing a condition or disease associated with the expression of SECP in a subject, comprising administering to said subject an effective amount of the composition of claim 33.

35. A method of preparing a polyclonal antibody with the specificity of the antibody of claim 10 comprising:

a) immunizing an animal with a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-44, or an immunogenic fragment thereof, under conditions to elicit an antibody response;

b) isolating antibodies from said animal; and

c) screening the isolated antibodies with the polypeptide, thereby identifying a polyclonal antibody which binds specifically to a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-44.

36. An antibody produced by a method of claim 35.

37. A composition comprising the antibody of claim 36 and a suitable carrier.

38. A method of making a monoclonal antibody with the specificity of the antibody of claim 10 comprising:

a) immunizing an animal with a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-44, or an immunogenic fragment thereof, under conditions to elicit an antibody response;

b) isolating antibody producing cells from the animal;

c) fusing the antibody producing cells with immortalized cells to form monoclonal antibody-producing hybridoma cells;

d) culturing the hybridoma cells; and

e) isolating from the culture monoclonal antibody which binds specifically to a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-44.

39. A monoclonal antibody produced by a method of claim 38.

40. A composition comprising the antibody of claim 39 and a suitable carrier.

41. The antibody of claim 10, wherein the antibody is produced by screening a Fab expression library.

42. The antibody of claim 10, wherein the antibody is produced by screening a recombinant immunoglobulin library.

43. A method for detecting a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-44 in a sample, comprising the steps of:

a) incubating the antibody of claim 10 with a sample under conditions to allow specific

binding of the antibody and the polypeptide; and

b) detecting specific binding, wherein specific binding indicates the presence of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-44 in the sample.

44. A method of purifying a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-44 from a sample, the method comprising:

a) incubating the antibody of claim 10 with a sample under conditions to allow specific binding of the antibody and the polypeptide; and

b) separating the antibody from the sample and obtaining the purified polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-44.

45. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:1.

46. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:2.

47. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:3.

48. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:4.

49. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:5.

50. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:6.

51. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:7.

52. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:8.

53. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:9.

54. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:10.

55. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:11.

56. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:12.

57. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:13.

58. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:14.

59. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:15.

60. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:16.

61. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:17.

62. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:18.

63. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:19.

64. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:20.

65. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:21.

66. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:22.

67. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:23.

68. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:24.

69. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:25.

70. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:26.

71. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:27.

72. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:28.

73. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:29.

74. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:30.

75. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:31.

76. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:32.

77. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:33.

78. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:34.

79. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:35.

80. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:36.

81. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:37.

82. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:38.

83. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:39.

84. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:40.

85. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:41.

86. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:42.

87. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:43.

88. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:44.

89. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:45.

90. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:46.

91. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:47.

92. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:48.

93. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:49.

94. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:50.

95. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:51.

96. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:52.

97. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:53.

98. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:54.

99. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:55.

100. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:56.

101. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:57.

102. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:58.

103. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:59.

104. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:60.

105. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:61.

106. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:62.

107. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:63.

108. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:64.

109. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:65.

110. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:66.

111. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:67.

112. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID

NO:68.

113. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:69.

114. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:70.

115. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:71.

116. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:72.

117. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:73.

118. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:74.

119. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:75.

120. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:76.

121. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:77.

122. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:78.

123. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:79.

124. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:80.

125. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:81.

126. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:82.

127. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:83.

128. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:84.

129. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:85.

130. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:86.

131. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:87.

132. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:88.

<110> INCYTE GENOMICS, INC.
HILLMAN, Jennifer L.
TANG, Y. Tom
YUE, Henry
ELLIOTT, Vicki S.
TRIBOULEY, Catherine M.
LEE, Ernestine A.
RAMKUMAR, Jayalaxmi
LAL, Preeti
XU, Yuming
WARREN, Bridget A.
HAFALIA, April J. A.
BAUGHN, Mariah R.
AZIMZAI, Yalda
BATRA, Sajeev
BURFORD, Neil
YAO, Monique G.
NGUYEN, Danniel B.
LU, Dyung Aina M.
WALIA, Narinder K.
GANDHI, Ameena R.
AU-YOUNG, Janice
PATTERSON, Chandra

<120> SECRETED PROTEINS

<130> PI-0133 PCT

<140> To Be Assigned
<141> Herewith

<150> 60/212,890; 60/213,466; 60/214,601; 60/222,372; 60/231,435; 60/232,889
<151> 2000-06-20; 2000-06-23; 2000-06-27; 2000-07-31; 2000-09-08; 2000-09-15

<160> 88
<170> PERL Program

<210> 1
<211> 552
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2101688CD1

<400> 1
```
Met Arg Arg Leu Thr Arg Arg Leu Val Leu Pro Val Phe Gly Val
  1               5                  10                  15
Leu Trp Ile Thr Val Leu Leu Phe Phe Trp Val Thr Lys Arg Lys
             20                  25                  30
Leu Glu Val Pro Thr Gly Pro Glu Val Gln Thr Pro Lys Pro Ser
             35                  40                  45
Asp Ala Asp Trp Asp Asp Leu Trp Asp Gln Phe Asp Glu Arg Arg
             50                  55                  60
Tyr Leu Asn Ala Lys Lys Trp Arg Val Gly Asp Asp Pro Tyr Lys
             65                  70                  75
Leu Tyr Ala Phe Asn Gln Arg Glu Ser Glu Arg Ile Ser Ser Asn
             80                  85                  90
Arg Ala Ile Pro Asp Thr Arg His Leu Arg Cys Thr Leu Leu Val
             95                 100                 105
Tyr Cys Thr Asp Leu Pro Pro Thr Ser Ile Ile Ile Thr Phe His
```

```
                110                 115                 120
Asn Glu Ala Arg Ser Thr Leu Leu Arg Thr Ile Arg Ser Val Leu
                125                 130                 135
Asn Arg Thr Pro Thr His Leu Ile Arg Glu Ile Ile Leu Val Asp
                140                 145                 150
Asp Phe Ser Asn Asp Pro Asp Asp Cys Lys Gln Leu Ile Lys Leu
                155                 160                 165
Pro Lys Val Lys Cys Leu Arg Asn Asn Glu Arg Gln Gly Leu Val
                170                 175                 180
Arg Ser Arg Ile Arg Gly Ala Asp Ile Ala Gln Gly Thr Thr Leu
                185                 190                 195
Thr Phe Leu Asp Ser His Cys Glu Val Asn Arg Asp Trp Leu Gln
                200                 205                 210
Pro Leu Leu His Arg Val Lys Glu Asp Tyr Thr Arg Val Val Cys
                215                 220                 225
Pro Val Ile Asp Ile Ile Asn Leu Asp Thr Phe Thr Tyr Ile Glu
                230                 235                 240
Ser Ala Ser Glu Leu Arg Gly Gly Phe Asp Trp Ser Leu His Phe
                245                 250                 255
Gln Trp Glu Gln Leu Ser Pro Glu Gln Lys Ala Arg Arg Leu Asp
                260                 265                 270
Pro Thr Glu Pro Ile Arg Thr Pro Ile Ile Ala Gly Gly Leu Phe
                275                 280                 285
Val Ile Asp Lys Ala Trp Phe Asp Tyr Leu Gly Lys Tyr Asp Met
                290                 295                 300
Asp Met Asp Ile Trp Gly Gly Glu Asn Phe Glu Ile Ser Phe Arg
                305                 310                 315
Val Trp Met Cys Gly Gly Ser Leu Glu Ile Val Pro Cys Ser Arg
                320                 325                 330
Val Gly His Val Phe Arg Lys Lys His Pro Tyr Val Phe Pro Asp
                335                 340                 345
Gly Asn Ala Asn Thr Tyr Ile Lys Asn Thr Lys Arg Thr Ala Glu
                350                 355                 360
Val Trp Met Asp Glu Tyr Lys Gln Tyr Tyr Tyr Ala Ala Arg Pro
                365                 370                 375
Phe Ala Leu Glu Arg Pro Phe Gly Asn Val Glu Ser Arg Leu Asp
                380                 385                 390
Leu Arg Lys Asn Leu Arg Cys Gln Ser Phe Lys Trp Tyr Leu Glu
                395                 400                 405
Asn Ile Tyr Pro Glu Leu Ser Ile Pro Lys Glu Ser Ser Ile Gln
                410                 415                 420
Lys Gly Asn Ile Arg Gln Arg Gln Lys Cys Leu Glu Ser Gln Arg
                425                 430                 435
Gln Asn Asn Gln Glu Thr Pro Asn Leu Lys Leu Ser Pro Cys Ala
                440                 445                 450
Lys Val Lys Gly Glu Asp Ala Lys Ser Gln Val Trp Ala Phe Thr
                455                 460                 465
Tyr Thr Gln Lys Ile Leu Gln Glu Glu Leu Cys Leu Ser Val Ile
                470                 475                 480
Thr Leu Phe Pro Gly Ala Pro Val Val Leu Val Leu Cys Lys Asn
                485                 490                 495
Gly Asp Asp Arg Gln Gln Trp Thr Lys Thr Gly Ser His Ile Glu
                500                 505                 510
His Ile Ala Ser His Leu Cys Leu Asp Thr Asp Met Phe Gly Asp
                515                 520                 525
Gly Thr Glu Asn Gly Lys Glu Ile Val Val Asn Pro Cys Glu Ser
                530                 535                 540
Ser Leu Met Ser Gln His Trp Asp Met Val Ser Ser
                545                 550
```

<210> 2
<211> 994
<212> PRT

<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 5452330CD1

<400> 2

```
Met Arg Pro Val Ala Leu Leu Leu Leu Pro Ser Leu Leu Ala Leu
  1               5                  10                  15
Leu Ala His Gly Leu Ser Leu Glu Ala Pro Thr Val Gly Lys Gly
             20                  25                  30
Gln Ala Pro Gly Ile Glu Glu Thr Asp Gly Glu Leu Thr Ala Ala
             35                  40                  45
Pro Thr Pro Glu Gln Pro Glu Arg Gly Val His Phe Val Thr Thr
             50                  55                  60
Ala Pro Thr Leu Lys Leu Leu Asn His His Pro Leu Leu Glu Glu
             65                  70                  75
Phe Leu Gln Glu Gly Leu Glu Lys Gly Asp Glu Glu Leu Arg Pro
             80                  85                  90
Ala Leu Pro Phe Gln Pro Asp Pro Pro Ala Pro Phe Thr Pro Ser
             95                 100                 105
Pro Leu Pro Arg Leu Ala Asn Gln Asp Ser Arg Pro Val Phe Thr
            110                 115                 120
Ser Pro Thr Pro Ala Met Ala Ala Val Pro Thr Gln Pro Gln Ser
            125                 130                 135
Lys Glu Gly Pro Trp Ser Pro Glu Ser Glu Ser Pro Met Leu Arg
            140                 145                 150
Ile Thr Ala Pro Leu Pro Pro Gly Pro Ser Met Ala Val Pro Thr
            155                 160                 165
Leu Gly Pro Gly Glu Ile Ala Ser Thr Thr Pro Pro Ser Arg Ala
            170                 175                 180
Trp Thr Pro Thr Gln Glu Gly Pro Gly Asp Met Gly Arg Pro Trp
            185                 190                 195
Val Ala Glu Val Val Ser Gln Gly Ala Gly Ile Gly Ile Gln Gly
            200                 205                 210
Thr Ile Thr Ser Ser Thr Ala Ser Gly Asp Asp Glu Glu Thr Thr
            215                 220                 225
Thr Thr Thr Thr Ile Ile Thr Thr Thr Ile Thr Thr Val Gln Thr
            230                 235                 240
Pro Gly Pro Cys Ser Trp Asn Phe Ser Gly Pro Glu Gly Ser Leu
            245                 250                 255
Asp Ser Pro Thr Asp Leu Ser Ser Pro Thr Asp Val Gly Leu Asp
            260                 265                 270
Cys Phe Phe Tyr Ile Ser Val Tyr Pro Gly Tyr Gly Val Glu Ile
            275                 280                 285
Lys Val Gln Asn Ile Ser Leu Arg Glu Gly Glu Thr Val Thr Val
            290                 295                 300
Glu Gly Leu Gly Gly Pro Asp Pro Leu Pro Leu Ala Asn Gln Ser
            305                 310                 315
Phe Leu Leu Arg Gly Gln Val Ile Arg Ser Pro Thr His Gln Ala
            320                 325                 330
Ala Leu Arg Phe Gln Ser Leu Pro Pro Pro Ala Gly Pro Gly Thr
            335                 340                 345
Phe His Phe His Tyr Gln Ala Tyr Leu Leu Ser Cys His Phe Pro
            350                 355                 360
Arg Arg Pro Ala Tyr Gly Asp Val Thr Val Thr Ser Leu His Pro
            365                 370                 375
Gly Gly Ser Ala Arg Phe His Cys Ala Thr Gly Tyr Gln Leu Lys
            380                 385                 390
Gly Ala Arg His Leu Thr Cys Leu Asn Ala Thr Gln Pro Phe Trp
            395                 400                 405
Asp Ser Lys Glu Pro Val Cys Ile Ala Ala Cys Gly Gly Val Ile
            410                 415                 420
```

```
Arg Asn Ala Thr Thr Gly Arg Ile Val Ser Pro Gly Phe Pro Gly
                425                 430                 435
Asn Tyr Ser Asn Asn Leu Thr Cys His Trp Leu Leu Glu Ala Pro
                440                 445                 450
Glu Gly Gln Arg Leu His Leu His Phe Glu Lys Val Ser Leu Ala
                455                 460                 465
Glu Asp Asp Asp Arg Leu Ile Ile Arg Asn Gly Asp Asn Val Glu
                470                 475                 480
Ala Pro Pro Val Tyr Asp Ser Tyr Glu Val Glu Tyr Leu Pro Ile
                485                 490                 495
Glu Gly Leu Leu Ser Ser Gly Lys His Phe Phe Val Glu Leu Ser
                500                 505                 510
Thr Asp Ser Ser Gly Ala Ala Ala Gly Met Ala Leu Arg Tyr Glu
                515                 520                 525
Ala Phe Gln Gln Gly His Cys Tyr Glu Pro Phe Val Lys Tyr Gly
                530                 535                 540
Asn Phe Ser Ser Ser Thr Pro Thr Tyr Pro Val Gly Thr Thr Val
                545                 550                 555
Glu Phe Ser Cys Asp Pro Gly Tyr Thr Leu Glu Gln Gly Ser Ile
                560                 565                 570
Ile Ile Glu Cys Val Asp Pro His Asp Pro Gln Trp Asn Glu Thr
                575                 580                 585
Glu Pro Ala Cys Arg Ala Val Cys Ser Gly Glu Ile Thr Asp Ser
                590                 595                 600
Ala Gly Val Val Leu Ser Pro Asn Trp Pro Glu Pro Tyr Gly Arg
                605                 610                 615
Gly Gln Asp Cys Ile Trp Gly Val His Val Glu Glu Asp Lys Arg
                620                 625                 630
Ile Met Leu Asp Ile Arg Val Leu Arg Ile Gly Pro Gly Asp Val
                635                 640                 645
Leu Thr Phe Tyr Asp Gly Asp Asp Leu Thr Ala Arg Val Leu Gly
                650                 655                 660
Gln Tyr Ser Gly Pro Arg Ser His Phe Lys Leu Phe Thr Ser Met
                665                 670                 675
Ala Asp Val Thr Ile Gln Phe Gln Ser Asp Pro Gly Thr Ser Val
                680                 685                 690
Leu Gly Tyr Gln Gln Gly Phe Val Ile His Phe Phe Glu Val Pro
                695                 700                 705
Arg Asn Asp Thr Cys Pro Glu Leu Pro Glu Ile Pro Asn Gly Trp
                710                 715                 720
Lys Ser Pro Ser Gln Pro Glu Leu Val His Gly Thr Val Val Thr
                725                 730                 735
Tyr Gln Cys Tyr Pro Gly Tyr Gln Val Val Gly Ser Ser Val Leu
                740                 745                 750
Met Cys Gln Trp Asp Leu Thr Trp Ser Glu Asp Leu Pro Ser Cys
                755                 760                 765
Gln Arg Val Thr Ser Cys His Asp Pro Gly Asp Val Glu His Ser
                770                 775                 780
Arg Arg Leu Ile Ser Ser Pro Lys Phe Pro Val Gly Ala Thr Val
                785                 790                 795
Gln Tyr Ile Cys Asp Gln Gly Phe Val Leu Met Gly Ser Ser Ile
                800                 805                 810
Leu Thr Cys His Asp Arg Gln Ala Gly Ser Pro Lys Trp Ser Asp
                815                 820                 825
Arg Ala Pro Lys Cys Leu Leu Glu Gln Leu Lys Pro Cys His Gly
                830                 835                 840
Leu Ser Ala Pro Glu Asn Gly Ala Arg Ser Pro Glu Lys Gln Leu
                845                 850                 855
His Pro Ala Gly Ala Thr Ile His Phe Ser Cys Ala Pro Gly Tyr
                860                 865                 870
Val Leu Lys Gly Gln Ala Ser Ile Lys Cys Val Pro Gly His Pro
                875                 880                 885
Ser His Trp Ser Asp Pro Pro Pro Ile Cys Arg Ala Ala Ser Leu
```

```
                890                 895                 900
Asp Gly Phe Tyr Asn Ser Arg Ser Leu Asp Val Ala Lys Ala Pro
                905                 910                 915
Ala Ala Ser Ser Thr Leu Asp Ala Ala His Ile Ala Ala Ala Ile
                920                 925                 930
Phe Leu Pro Leu Val Ala Met Val Leu Leu Val Gly Gly Val Tyr
                935                 940                 945
Phe Tyr Phe Ser Arg Leu Gln Gly Lys Ser Ser Leu Gln Leu Pro
                950                 955                 960
Arg Pro Arg Pro Arg Pro Tyr Asn Arg Ile Thr Ile Glu Ser Ala
                965                 970                 975
Phe Asp Asn Pro Thr Tyr Glu Thr Gly Ser Leu Ser Phe Ala Gly
                980                 985                 990
Asp Glu Arg Ile
```

<210> 3
<211> 212
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 4362432CD1

<400> 3

```
Met Leu Ser Ser Val Val Phe Trp Gly Leu Ile Ala Leu Ile Gly
  1                 5                  10                  15
Thr Ser Arg Gly Ser Tyr Pro Phe Ser His Ser Met Lys Pro His
                 20                  25                  30
Leu His Pro Arg Leu Tyr His Gly Cys Tyr Gly Asp Ile Met Thr
                 35                  40                  45
Met Lys Thr Ser Gly Ala Thr Cys Asp Ala Asn Ser Val Met Asn
                 50                  55                  60
Cys Gly Ile Arg Gly Ser Glu Met Phe Ala Glu Met Asp Leu Arg
                 65                  70                  75
Ala Ile Lys Pro Tyr Gln Thr Leu Ile Lys Glu Val Gly Gln Arg
                 80                  85                  90
His Cys Val Asp Pro Ala Val Ile Ala Ala Ile Ile Ser Arg Glu
                 95                 100                 105
Ser His Gly Gly Ser Val Leu Gln Asp Gly Trp Asp His Arg Gly
                110                 115                 120
Leu Lys Phe Gly Leu Met Gln Leu Asp Lys Gln Thr Tyr His Pro
                125                 130                 135
Val Gly Ala Trp Asp Ser Lys Glu His Leu Ser Gln Ala Thr Gly
                140                 145                 150
Ile Leu Thr Glu Arg Ile Lys Ala Ile Gln Lys Lys Phe Pro Thr
                155                 160                 165
Trp Ser Val Ala Gln His Leu Lys Gly Gly Leu Ser Ala Phe Lys
                170                 175                 180
Ser Gly Ile Glu Ala Ile Ala Thr Pro Ser Asp Ile Asp Asn Asp
                185                 190                 195
Phe Val Asn Asp Ile Ile Ala Arg Ala Lys Phe Tyr Lys Arg Gln
                200                 205                 210
Ser Phe
```

<210> 4
<211> 308
<212> PRT
<213> Homo sapiens

<220>

<221> misc_feature
<223> Incyte ID No: 5308104CD1

<400> 4
Met Asn Gly Leu Ser Leu Ser Glu Leu Cys Cys Leu Phe Cys Cys
1 5 10 15
Pro Pro Cys Pro Gly Arg Ile Ala Ala Lys Leu Ala Phe Leu Pro
20 25 30
Pro Glu Ala Thr Tyr Ser Leu Val Pro Glu Pro Glu Pro Gly Pro
35 40 45
Gly Gly Ala Gly Ala Ala Pro Leu Gly Thr Leu Arg Ala Ser Ser
50 55 60
Gly Ala Pro Gly Arg Trp Lys Leu His Leu Thr Glu Arg Ala Asp
65 70 75
Phe Gln Tyr Ser Gln Arg Glu Leu Asp Thr Ile Glu Val Phe Pro
80 85 90
Thr Lys Ser Ala Arg Gly Asn Arg Val Ser Cys Met Tyr Val Arg
95 100 105
Cys Val Pro Gly Ala Arg Tyr Thr Val Leu Phe Ser His Gly Asn
110 115 120
Ala Val Asp Leu Gly Gln Met Ser Ser Phe Tyr Ile Gly Leu Gly
125 130 135
Ser Arg Leu His Cys Asn Ile Phe Ser Tyr Asp Tyr Ser Gly Tyr
140 145 150
Gly Ala Ser Ser Gly Arg Pro Ser Glu Arg Asn Leu Tyr Ala Asp
155 160 165
Ile Asp Ala Ala Trp Gln Ala Leu Arg Thr Arg Tyr Gly Ile Ser
170 175 180
Pro Asp Ser Ile Ile Leu Tyr Gly Gln Ser Ile Gly Thr Val Pro
185 190 195
Thr Val Asp Leu Ala Ser Arg Tyr Glu Cys Ala Ala Val Ile Leu
200 205 210
His Ser Pro Leu Met Ser Gly Leu Arg Val Ala Phe Pro Asp Thr
215 220 225
Arg Lys Thr Tyr Cys Phe Asp Ala Phe Pro Ser Ile Asp Lys Ile
230 235 240
Ser Lys Val Thr Ser Pro Val Leu Val Ile His Gly Thr Glu Asp
245 250 255
Glu Val Ile Asp Phe Ser His Gly Leu Ala Met Tyr Glu Arg Cys
260 265 270
Pro Arg Ala Val Glu Pro Leu Trp Val Glu Gly Ala Gly His Asn
275 280 285
Asp Ile Glu Leu Tyr Ala Gln Tyr Leu Glu Arg Leu Lys Gln Phe
290 295 300
Ile Ser His Glu Leu Pro Asn Ser
305

<210> 5
<211> 328
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3092736CD1

<400> 5
Met Trp Cys Ser Phe Leu Ala Pro Val Ser Ser Ser Cys Leu Cys
1 5 10 15
Trp Val Trp Ala Cys Trp Gly Glu Gly His Cys Cys Gln Arg Gly
20 25 30
Thr Asp Phe Leu Met Val Leu Pro Lys Val Asn Val Gly Asp Thr
35 40 45

```
Val Ala Met Leu Pro Lys Ser Arg Arg Ala Leu Thr Ile Gln Glu
                50                  55                  60
Ile Ala Ala Leu Ala Arg Ser Ser Leu His Gly Ile Ser Gln Val
                65                  70                  75
Val Lys Asp His Val Thr Lys Pro Thr Ala Met Ala Gln Gly Arg
                80                  85                  90
Val Ala His Leu Ile Glu Trp Lys Gly Trp Ser Lys Pro Ser Asp
                95                 100                 105
Ser Pro Ala Ala Leu Glu Ser Ala Phe Ser Ser Tyr Ser Asp Leu
               110                 115                 120
Ser Glu Gly Glu Gln Glu Ala Arg Phe Ala Ala Gly Val Ala Glu
               125                 130                 135
Gln Phe Ala Ile Ala Glu Ala Lys Leu Arg Ala Trp Ser Ser Val
               140                 145                 150
Asp Gly Glu Asp Ser Thr Asp Asp Ser Tyr Asp Glu Asp Phe Ala
               155                 160                 165
Gly Gly Met Asp Thr Asp Met Ala Gly Gln Leu Pro Leu Gly Pro
               170                 175                 180
His Leu Gln Asp Leu Phe Thr Gly His Arg Phe Ser Arg Pro Val
               185                 190                 195
Arg Gln Gly Ser Val Glu Pro Glu Ser Asp Cys Ser Gln Thr Val
               200                 205                 210
Ser Pro Asp Thr Leu Cys Ser Ser Leu Cys Ser Leu Glu Asp Gly
               215                 220                 225
Leu Leu Gly Ser Pro Ala Arg Leu Ala Ser Gln Leu Leu Gly Asp
               230                 235                 240
Glu Leu Leu Leu Ala Lys Leu Pro Pro Ser Arg Glu Ser Ala Phe
               245                 250                 255
Arg Ser Leu Gly Pro Leu Glu Ala Gln Asp Ser Leu Tyr Asn Ser
               260                 265                 270
Pro Leu Thr Glu Ser Cys Leu Ser Pro Ala Glu Glu Glu Pro Ala
               275                 280                 285
Pro Cys Lys Asp Cys Gln Pro Leu Cys Pro Pro Leu Thr Gly Ser
               290                 295                 300
Trp Glu Arg Gln Arg Gln Ala Ser Asp Leu Ala Ser Ser Gly Val
               305                 310                 315
Val Ser Leu Asp Glu Asp Glu Ala Glu Pro Glu Glu Gln
               320                 325

<210> 6
<211> 69
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3580257CD1

<400> 6
Met Ala Met Ala Val Asp Val Ala Ala Ser Ala Asp Gly Val Leu
  1                 5                  10                  15
Ala Val Ala Met Glu Ala Thr Asp Met Ala Leu Ala Leu Glu Ala
                20                  25                  30
Thr Asp Met Ala Leu Ala Leu Glu Ala Thr Asp Met Ala Leu Ala
                35                  40                  45
Leu Glu Ala Met Asp Met Ala Ala Ala Ala His Arg Thr Met Glu
                50                  55                  60
Asp Thr Asp Ser Leu Ala Phe Ile Lys
                65

<210> 7
<211> 158
<212> PRT
```

<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3634758CD1

<400> 7
Met Ala Asp Glu Ala Leu Phe Leu Leu Leu His Asn Glu Met Val
1 5 10 15
Ser Gly Val Tyr Lys Ser Ala Glu Gln Gly Glu Val Glu Asn Gly
20 25 30
Arg Cys Ile Thr Lys Leu Glu Asn Met Gly Phe Arg Val Gly Gln
35 40 45
Gly Leu Ile Glu Arg Phe Thr Lys Asp Thr Ala Arg Phe Lys Asp
50 55 60
Glu Leu Asp Ile Met Lys Phe Ile Cys Lys Asp Phe Trp Thr Thr
65 70 75
Val Phe Lys Lys Gln Ile Asp Asn Leu Arg Thr Asn His Gln Gly
80 85 90
Ile Tyr Val Leu Gln Asp Asn Lys Phe Arg Leu Leu Thr Gln Met
95 100 105
Ser Ala Gly Lys Gln Tyr Leu Glu His Ala Ser Lys Tyr Leu Ala
110 115 120
Phe Thr Cys Gly Leu Ile Arg Gly Gly Leu Ser Asn Leu Gly Ile
125 130 135
Lys Ser Ile Val Thr Ala Glu Val Ser Ser Met Pro Ala Cys Lys
140 145 150
Phe Gln Val Met Ile Gln Lys Leu
155

<210> 8
<211> 463
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 4027923CD1

<400> 8
Met Arg Ala Gly Pro Glu Pro Gln Ala Leu Val Gly Gln Lys Arg
1 5 10 15
Gly Ala Leu Arg Leu Leu Val Pro Arg Leu Val Leu Thr Val Ser
20 25 30
Ala Pro Ala Glu Val Arg Arg Arg Val Leu Arg Pro Val Leu Ser
35 40 45
Trp Met Asp Arg Glu Thr Arg Ala Leu Ala Asp Ser His Phe Arg
50 55 60
Gly Leu Gly Val Asp Val Pro Gly Val Gly Gln Ala Pro Gly Arg
65 70 75
Val Ala Phe Val Ser Glu Pro Gly Ala Phe Ser Tyr Ala Asp Phe
80 85 90
Val Arg Gly Phe Leu Leu Pro Asn Leu Pro Cys Val Phe Ser Ser
95 100 105
Ala Phe Thr Gln Gly Trp Gly Ser Arg Arg Arg Trp Val Thr Pro
110 115 120
Ala Gly Arg Pro Asp Phe Asp His Leu Leu Arg Thr Tyr Gly Asp
125 130 135
Val Val Val Pro Val Ala Asn Cys Gly Val Gln Glu Tyr Asn Ser
140 145 150
Asn Pro Lys Glu His Met Thr Leu Arg Asp Tyr Ile Thr Tyr Trp
155 160 165
Lys Glu Tyr Ile Gln Ala Gly Tyr Ser Ser Pro Arg Gly Cys Leu

```
                170                 175                 180
Tyr Leu Lys Asp Trp His Leu Cys Arg Asp Phe Pro Val Glu Asp
                185                 190                 195
Val Phe Thr Leu Pro Val Tyr Phe Ser Ser Asp Trp Leu Asn Glu
                200                 205                 210
Phe Trp Asp Ala Leu Asp Val Asp Asp Tyr Arg Phe Val Tyr Ala
                215                 220                 225
Gly Pro Ala Gly Ser Trp Ser Pro Phe His Ala Asp Ile Phe Arg
                230                 235                 240
Ser Phe Ser Trp Ser Val Asn Val Cys Gly Arg Lys Lys Trp Leu
                245                 250                 255
Leu Phe Pro Pro Gly Gln Glu Glu Ala Leu Arg Asp Arg His Gly
                260                 265                 270
Asn Leu Pro Tyr Asp Val Thr Ser Pro Ala Leu Cys Asp Thr His
                275                 280                 285
Leu His Pro Arg Asn Gln Leu Ala Gly Pro Pro Leu Glu Ile Thr
                290                 295                 300
Gln Glu Ala Gly Glu Met Val Phe Val Pro Ser Gly Trp His His
                305                 310                 315
Gln Val His Asn Leu Asp Asp Thr Ile Ser Ile Asn His Asn Trp
                320                 325                 330
Val Asn Gly Phe Asn Leu Ala Asn Met Trp Arg Phe Leu Gln Gln
                335                 340                 345
Glu Leu Cys Ala Val Gln Glu Glu Val Ser Glu Trp Arg Asp Ser
                350                 355                 360
Met Pro Asp Trp His His His Cys Gln Val Ile Met Arg Ser Cys
                365                 370                 375
Ser Gly Ile Asn Phe Glu Glu Phe Tyr His Phe Leu Lys Val Ile
                380                 385                 390
Ala Glu Lys Arg Leu Leu Val Leu Arg Glu Ala Ala Ala Glu Asp
                395                 400                 405
Gly Ala Gly Leu Gly Phe Glu Gln Ala Ala Phe Asp Val Gly Arg
                410                 415                 420
Ile Thr Glu Val Leu Ala Ser Leu Val Ala His Pro Asp Phe Gln
                425                 430                 435
Arg Val Asp Thr Ser Ala Phe Ser Pro Gln Pro Lys Glu Leu Leu
                440                 445                 450
Gln Gln Leu Arg Glu Ala Val Asp Ala Ala Ala Ala Pro
                455                 460
```

<210> 9
<211> 648
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 4348533CD1

<400> 9

```
Met Glu Lys Ala Arg Arg Gly Gly Asp Gly Val Pro Arg Gly Pro
  1                 5                  10                  15
Val Leu His Ile Val Val Val Gly Phe His His Lys Lys Gly Cys
                 20                  25                  30
Gln Val Glu Phe Ser Tyr Pro Pro Leu Ile Pro Gly Asp Gly His
                 35                  40                  45
Asp Ser His Thr Leu Pro Glu Glu Trp Lys Tyr Leu Pro Phe Leu
                 50                  55                  60
Ala Leu Pro Asp Gly Ala His Asn Tyr Gln Glu Asp Thr Val Phe
                 65                  70                  75
Phe His Leu Pro Pro Arg Asn Gly Asn Gly Ala Thr Val Phe Gly
                 80                  85                  90
Ile Ser Cys Tyr Arg Gln Ile Glu Ala Lys Ala Leu Lys Val Arg
```

```
                95                  100                 105
Gln Ala Asp Ile Thr Arg Glu Thr Val Gln Lys Ser Val Cys Val
               110                 115                 120
Leu Ser Lys Leu Pro Leu Tyr Gly Leu Leu Gln Ala Lys Leu Gln
               125                 130                 135
Leu Ile Thr His Ala Tyr Phe Glu Glu Lys Asp Phe Ser Gln Ile
               140                 145                 150
Ser Ile Leu Lys Glu Leu Tyr Glu His Met Asn Ser Ser Leu Gly
               155                 160                 165
Gly Ala Ser Leu Glu Gly Ser Gln Val Tyr Leu Gly Leu Ser Pro
               170                 175                 180
Arg Asp Leu Val Leu His Phe Arg His Lys Val Leu Ile Leu Phe
               185                 190                 195
Lys Leu Ile Leu Leu Glu Lys Lys Val Leu Phe Tyr Ile Ser Pro
               200                 205                 210
Val Asn Lys Leu Val Gly Ala Leu Met Thr Val Leu Ser Leu Phe
               215                 220                 225
Pro Gly Met Ile Glu His Gly Leu Ser Asp Cys Ser Gln Tyr Arg
               230                 235                 240
Pro Arg Lys Ser Met Ser Glu Asp Gly Gly Leu Gln Glu Ser Asn
               245                 250                 255
Pro Cys Ala Asp Asp Phe Val Ser Ala Ser Thr Ala Asp Val Ser
               260                 265                 270
His Thr Asn Leu Gly Thr Ile Arg Lys Val Met Ala Gly Asn His
               275                 280                 285
Gly Glu Asp Ala Ala Met Lys Thr Glu Glu Pro Leu Phe Gln Val
               290                 295                 300
Glu Asp Ser Ser Lys Gly Gln Glu Pro Asn Asp Thr Asn Gln Tyr
               305                 310                 315
Leu Lys Pro Pro Ser Arg Pro Ser Pro Asp Ser Ser Glu Ser Asp
               320                 325                 330
Trp Glu Thr Leu Asp Pro Ser Val Leu Glu Asp Pro Asn Leu Lys
               335                 340                 345
Glu Arg Glu Gln Leu Gly Ser Asp Gln Thr Asn Leu Phe Pro Lys
               350                 355                 360
Asp Ser Val Pro Ser Glu Ser Leu Pro Ile Thr Val Gln Pro Gln
               365                 370                 375
Ala Asn Thr Gly Gln Val Val Leu Ile Pro Gly Leu Ile Ser Gly
               380                 385                 390
Leu Glu Glu Asp Gln Tyr Gly Met Pro Leu Ala Ile Phe Thr Lys
               395                 400                 405
Gly Tyr Leu Cys Leu Pro Tyr Met Ala Leu Gln Gln His His Leu
               410                 415                 420
Leu Ser Asp Val Thr Val Arg Gly Phe Val Ala Gly Ala Thr Asn
               425                 430                 435
Ile Leu Phe Arg Gln Gln Lys His Leu Ser Asp Ala Ile Val Glu
               440                 445                 450
Val Glu Glu Ala Leu Ile Gln Ile His Asp Pro Glu Leu Arg Lys
               455                 460                 465
Leu Leu Asn Pro Thr Thr Ala Asp Leu Arg Phe Ala Asp Tyr Leu
               470                 475                 480
Val Arg His Val Thr Glu Asn Arg Asp Asp Val Phe Leu Asp Gly
               485                 490                 495
Thr Gly Trp Glu Gly Gly Asp Glu Trp Ile Arg Ala Gln Phe Ala
               500                 505                 510
Val Tyr Ile His Ala Leu Leu Ala Ala Thr Leu Gln Leu Asp Asn
               515                 520                 525
Glu Lys Ile Leu Ser Asp Tyr Gly Thr Thr Phe Val Thr Ala Trp
               530                 535                 540
Lys Asn Thr His Asn Tyr Arg Val Trp Asn Ser Asn Lys His Pro
               545                 550                 555
Ala Leu Ala Glu Ile Asn Pro Asn His Pro Phe Gln Gly Gln Tyr
               560                 565                 570
```

```
Ser Val Ser Asp Met Lys Leu Arg Phe Ser His Ser Val Gln Asn
                575                 580                 585
Ser Glu Arg Gly Lys Lys Ile Gly Asn Val Met Val Thr Thr Ser
                590                 595                 600
Arg Asn Val Val Gln Thr Gly Lys Ala Val Gly Gln Ser Val Gly
                605                 610                 615
Gly Ala Phe Ser Ser Ala Lys Thr Ala Met Ser Ser Trp Leu Ser
                620                 625                 630
Thr Phe Thr Thr Ser Thr Ser Gln Ser Leu Thr Glu Pro Pro Asp
                635                 640                 645
Glu Lys Pro
```

<210> 10
<211> 130
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 4521857CD1

<400> 10

```
Met Tyr Leu Gln Val Glu Thr Arg Thr Ser Ser Arg Leu His Leu
  1               5                  10                  15
Lys Arg Ala Pro Gly Ile Arg Ser Trp Ser Leu Leu Val Gly Ile
                20                  25                  30
Leu Ser Ile Gly Leu Ala Ala Ala Tyr Tyr Ser Gly Asp Ser Leu
                35                  40                  45
Gly Trp Lys Leu Phe Tyr Val Thr Gly Cys Leu Phe Val Ala Val
                50                  55                  60
Gln Asn Leu Glu Asp Trp Glu Glu Ala Ile Phe Asp Lys Ser Thr
                65                  70                  75
Gly Lys Val Val Leu Lys Thr Phe Ser Leu Tyr Lys Lys Leu Leu
                80                  85                  90
Thr Leu Phe Arg Ala Gly His Asp Gln Val Val Val Leu Leu His
                95                 100                 105
Val Val Pro Asp Thr Ala Ser Ser Pro Trp Trp Thr Ser Pro Ala
               110                 115                 120
Val Arg Cys Phe Pro Lys Gly Ser Glu Gly
               125                 130
```

<210> 11
<211> 279
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 4722253CD1

<400> 11

```
Met Gly Arg Gly Leu Arg Trp Trp Gly Gly Arg Gly Arg Arg His
  1               5                  10                  15
Gly Gln Ala Pro Glu Trp Gly Pro Leu Val Gly Ala Arg Leu Lys
                20                  25                  30
Gly Val Ala Arg Ala Ala Ser Leu Val Gly Arg Arg Arg Ala Gly
                35                  40                  45
Thr Gly Met Ala Leu Leu Leu Cys Leu Val Cys Leu Thr Ala Ala
                50                  55                  60
Leu Ala His Gly Cys Leu His Cys His Ser Asn Phe Ser Lys Lys
                65                  70                  75
Phe Ser Phe Tyr Arg His His Val Asn Phe Lys Ser Trp Trp Val
```

```
                 80                  85                  90
Gly Asp Ile Pro Val Ser Gly Ala Leu Leu Thr Asp Trp Ser Asp
                 95                 100                 105
Asp Thr Met Lys Glu Leu His Leu Ala Ile Pro Ala Lys Ile Thr
                110                 115                 120
Arg Glu Lys Leu Asp Gln Val Ala Thr Ala Val Tyr Gln Met Met
                125                 130                 135
Asp Gln Leu Tyr Gln Gly Lys Met Tyr Phe Pro Gly Tyr Phe Pro
                140                 145                 150
Asn Glu Leu Arg Asn Ile Phe Arg Glu Gln Val His Leu Ile Gln
                155                 160                 165
Asn Ala Ile Ile Glu Ser Arg Ile Asp Cys Gln His Arg Cys Gly
                170                 175                 180
Ile Phe Gln Tyr Glu Thr Ile Ser Cys Asn Asn Cys Thr Asp Ser
                185                 190                 195
His Val Ala Cys Phe Gly Tyr Asn Cys Glu Ser Ser Ala Gln Trp
                200                 205                 210
Lys Ser Ala Val Gln Gly Leu Leu Asn Tyr Ile Asn Asn Trp His
                215                 220                 225
Lys Gln Asp Thr Ser Met Arg Pro Arg Ser Ser Ala Phe Ser Trp
                230                 235                 240
Pro Gly Thr His Arg Ala Thr Pro Ala Phe Leu Val Ser Pro Ala
                245                 250                 255
Leu Arg Cys Leu Glu Pro Pro His Leu Ala Asn Leu Thr Leu Glu
                260                 265                 270
Asp Ala Ala Glu Cys Leu Lys Gln His
                275
```

<210> 12
<211> 458
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 4878134CD1

<400> 12

```
Met Pro Thr Ile Leu Trp Leu Met Asp Trp Ser Asp Met Asn Ser
  1                 5                  10                  15
Asn Leu Asp Leu Leu Ala Leu Leu Gly Leu Gly Ile Ser Ser Phe
                 20                  25                  30
Val Leu Ile Thr Gly Cys Ala Asn Met Leu Leu Met Ala Ala Leu
                 35                  40                  45
Trp Gly Leu Tyr Met Ser Leu Val Asn Val Gly His Val Trp Tyr
                 50                  55                  60
Ser Phe Gly Trp Glu Ser Gln Leu Leu Glu Thr Gly Phe Leu Gly
                 65                  70                  75
Ile Phe Leu Cys Pro Leu Trp Thr Leu Ser Arg Leu Pro Gln His
                 80                  85                  90
Thr Pro Thr Ser Arg Ile Val Leu Trp Gly Phe Arg Trp Leu Ile
                 95                 100                 105
Phe Arg Ile Met Leu Gly Ala Gly Leu Ile Lys Ile Arg Gly Asp
                110                 115                 120
Arg Cys Trp Arg Asp Leu Thr Cys Met Asp Phe His Tyr Glu Thr
                125                 130                 135
Gln Pro Met Pro Asn Pro Val Ala Tyr Tyr Leu His His Ser Pro
                140                 145                 150
Trp Trp Phe His Arg Phe Glu Thr Leu Ser Asn His Phe Ile Glu
                155                 160                 165
Leu Leu Val Pro Phe Phe Leu Phe Leu Gly Arg Arg Ala Cys Ile
                170                 175                 180
Ile His Gly Val Leu Gln Ile Leu Phe Gln Ala Val Leu Ile Val
```

```
                185                 190                 195
Ser Gly Asn Leu Ser Phe Leu Asn Trp Leu Thr Met Val Pro Ser
                200                 205                 210
Leu Ala Cys Phe Asp Asp Ala Thr Leu Gly Phe Leu Phe Pro Ser
                215                 220                 225
Gly Pro Gly Ser Leu Lys Asp Arg Val Leu Gln Met Gln Arg Asp
                230                 235                 240
Ile Arg Gly Ala Arg Pro Glu Pro Arg Phe Gly Ser Val Val Arg
                245                 250                 255
Arg Ala Ala Asn Val Ser Leu Gly Val Leu Leu Ala Trp Leu Ser
                260                 265                 270
Val Pro Val Val Leu Asn Leu Leu Ser Ser Arg Gln Val Met Asn
                275                 280                 285
Thr His Phe Asn Ser Leu His Ile Val Asn Thr Tyr Gly Ala Phe
                290                 295                 300
Gly Ser Ile Thr Lys Glu Arg Ala Glu Val Ile Leu Gln Gly Thr
                305                 310                 315
Ala Ser Ser Asn Ala Ser Ala Pro Asp Ala Met Trp Glu Asp Tyr
                320                 325                 330
Glu Phe Lys Cys Lys Pro Gly Asp Pro Ser Arg Arg Pro Cys Leu
                335                 340                 345
Ile Ser Pro Tyr His Tyr Arg Leu Asp Trp Leu Met Trp Phe Ala
                350                 355                 360
Ala Phe Gln Thr Tyr Glu His Asn Asp Trp Ile Ile His Leu Ala
                365                 370                 375
Gly Lys Leu Leu Ala Ser Asp Ala Glu Ala Leu Ser Leu Leu Ala
                380                 385                 390
His Asn Pro Phe Ala Gly Arg Pro Pro Pro Arg Trp Val Arg Gly
                395                 400                 405
Glu His Tyr Arg Tyr Lys Phe Ser Arg Pro Gly Gly Arg His Ala
                410                 415                 420
Ala Glu Gly Lys Trp Trp Val Arg Lys Arg Ile Gly Ala Tyr Phe
                425                 430                 435
Pro Pro Leu Ser Leu Glu Glu Leu Arg Pro Tyr Phe Arg Asp Arg
                440                 445                 450
Gly Trp Pro Leu Pro Gly Pro Leu
                455

<210> 13
<211> 173
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 5050133CD1

<400> 13
Met Leu Leu Val Asp Ala Asp Gln Pro Glu Pro Met Arg Ser Gly
  1               5                  10                  15
Ala Arg Glu Leu Ala Leu Phe Leu Thr Pro Glu Pro Gly Ala Glu
                 20                  25                  30
Ala Lys Glu Val Glu Glu Thr Ile Glu Gly Met Leu Leu Arg Leu
                 35                  40                  45
Glu Glu Phe Cys Ser Leu Ala Asp Leu Ile Arg Ser Asp Thr Ser
                 50                  55                  60
Gln Ile Leu Glu Glu Asn Ile Pro Val Leu Lys Ala Lys Leu Thr
                 65                  70                  75
Glu Met Arg Gly Ile Tyr Ala Lys Val Asp Arg Leu Glu Ala Phe
                 80                  85                  90
Val Lys Met Val Gly His His Val Ala Phe Leu Glu Ala Asp Val
                 95                 100                 105
Leu Gln Ala Glu Arg Asp His Gly Ala Phe Pro Gln Ala Leu Arg
```

```
                110                 115                 120
Arg Trp Leu Gly Ser Ala Gly Leu Pro Ser Phe Arg Asn Lys Ser
                125                 130                 135
Pro Ala Pro Val Pro Val Thr Tyr Glu Leu Pro Thr Leu Tyr Arg
                140                 145                 150
Thr Glu Asp Tyr Phe Pro Val Asp Ala Gly Glu Ala Gln His His
                155                 160                 165
Pro Arg Thr Cys Pro Arg Pro Leu
                170
```

<210> 14
<211> 335
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 5630124CD1

<400> 14

```
Met Gly Ala Ser Ser Ser Ser Ala Leu Ala Arg Leu Gly Leu Pro
  1                 5                  10                  15
Ala Arg Pro Trp Pro Arg Trp Leu Gly Val Ala Ala Leu Gly Leu
                 20                  25                  30
Ala Ala Val Ala Leu Gly Thr Val Ala Trp Arg Arg Ala Trp Pro
                 35                  40                  45
Arg Arg Arg Arg Arg Leu Gln Gln Val Gly Thr Val Ala Lys Leu
                 50                  55                  60
Trp Ile Tyr Pro Val Lys Ser Cys Lys Gly Val Pro Val Ser Glu
                 65                  70                  75
Ala Glu Cys Thr Ala Met Gly Leu Arg Ser Gly Asn Leu Arg Asp
                 80                  85                  90
Arg Phe Trp Leu Val Ile Lys Glu Asp Gly His Met Val Thr Ala
                 95                 100                 105
Arg Gln Glu Pro Arg Leu Val Leu Ile Ser Ile Ile Tyr Glu Asn
                110                 115                 120
Asn Cys Leu Ile Phe Arg Ala Pro Asp Met Asp Gln Leu Val Leu
                125                 130                 135
Pro Ser Lys Gln Pro Ser Ser Asn Lys Leu His Asn Cys Arg Ile
                140                 145                 150
Phe Gly Leu Asp Ile Lys Gly Arg Asp Cys Gly Asn Glu Ala Ala
                155                 160                 165
Lys Trp Phe Thr Asn Phe Leu Lys Thr Glu Ala Tyr Arg Leu Val
                170                 175                 180
Gln Phe Glu Thr Asn Met Lys Gly Arg Thr Ser Arg Lys Leu Leu
                185                 190                 195
Pro Thr Leu Asp Gln Asn Phe Gln Val Ala Tyr Pro Asp Tyr Cys
                200                 205                 210
Pro Leu Leu Ile Met Thr Asp Ala Ser Leu Val Asp Leu Asn Thr
                215                 220                 225
Arg Met Glu Lys Lys Met Lys Met Glu Asn Phe Arg Pro Asn Ile
                230                 235                 240
Val Val Thr Gly Cys Asp Ala Phe Glu Glu Asp Thr Trp Asp Glu
                245                 250                 255
Leu Leu Ile Gly Ser Val Glu Val Lys Lys Val Met Ala Cys Pro
                260                 265                 270
Arg Cys Ile Leu Thr Thr Val Asp Pro Asp Thr Gly Val Ile Asp
                275                 280                 285
Arg Lys Gln Pro Leu Asp Thr Leu Lys Ser Tyr Arg Leu Cys Asp
                290                 295                 300
Pro Ser Glu Arg Glu Leu Tyr Lys Leu Ser Pro Leu Phe Gly Ile
                305                 310                 315
Tyr Tyr Ser Val Glu Lys Ile Gly Ser Leu Arg Val Gly Asp Pro
```

320 325 330
Val Tyr Arg Met Val
335

<210> 15
<211> 71
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 5677286CD1

<400> 15
Met His Ser Pro Ala Ser Gly Pro Leu Leu Pro Pro Leu Arg Val
1 5 10 15
Pro Trp Leu Pro Pro Val Val Leu Gly Asn Leu Gly Pro Ser Pro
20 25 30
Ala Ser Pro Ala Ser His Ser Ser Ser Leu Val Thr Leu Arg Glu
35 40 45
Leu Arg Ala Arg Leu Val Ala Gly Leu Leu Cys Phe Cys Pro Arg
50 55 60
Leu Leu Trp Ser Leu Ala Gly Asn Ser Met Ile
65 70

<210> 16
<211> 148
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 6436791CD1

<400> 16
Met Leu Pro Arg Gly Leu Lys Met Ala Pro Arg Gly Lys Arg Leu
1 5 10 15
Ser Ser Thr Pro Leu Glu Ile Leu Phe Phe Leu Asn Gly Trp Tyr
20 25 30
Asn Ala Thr Tyr Phe Leu Leu Glu Leu Phe Ile Phe Leu Tyr Lys
35 40 45
Gly Val Leu Leu Pro Tyr Pro Thr Ala Asn Leu Val Leu Asp Val
50 55 60
Val Met Leu Leu Leu Tyr Leu Gly Ile Glu Val Ile Arg Leu Phe
65 70 75
Phe Gly Thr Lys Gly Asn Leu Cys Gln Arg Lys Met Pro Leu Ser
80 85 90
Ile Ser Val Ala Leu Thr Phe Pro Ser Ala Met Met Ala Ser Tyr
95 100 105
Tyr Leu Leu Leu Gln Thr Tyr Val Leu Arg Leu Glu Ala Ile Met
110 115 120
Asn Gly Ile Leu Leu Phe Phe Cys Gly Ser Glu Leu Leu Leu Glu
125 130 135
Val Leu Thr Leu Ala Ala Phe Ser Ser Met Asp Thr Ile
140 145

<210> 17
<211> 231
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature

<223> Incyte ID No: 1820972CD1

<400> 17
Met Ala Trp Ile Pro Leu Phe Leu Gly Val Leu Ala Tyr Cys Thr
1 5 10 15
Gly Ser Met Asp Ser Phe Glu Leu Thr Gln Ala Pro Ser Thr Ser
20 25 30
Val Ser Pro Gly Gln Thr Ala Thr Ile Ser Cys Ser Gly Glu Lys
35 40 45
Val Gly Ser Lys Phe Phe Ser Trp Tyr Gln Gln Lys Glu Gly Gln
50 55 60
Ser Pro Val Val Ile Ile Tyr Gln Asn Gly Lys Arg Pro Ser Glu
65 70 75
Ile Ala Asp Arg Phe Ser Gly Ser Lys Ser Gly Asp Thr Ala Thr
80 85 90
Leu Thr Ile Ser Arg Ala Gln Ala Gly Asp Glu Ala Asp Tyr Phe
95 100 105
Cys Gln Val Trp Asp Ser Ser Thr Ala Val Phe Gly Gly Gly Thr
110 115 120
Lys Leu Thr Val Leu Gly Gln Pro Lys Ala Ala Pro Ser Val Thr
125 130 135
Leu Phe Pro Pro Ser Ser Glu Glu Leu Gln Ala Asn Lys Ala Thr
140 145 150
Leu Val Cys Leu Ile Ser Asp Phe Tyr Pro Gly Ala Val Thr Val
155 160 165
Ala Trp Lys Ala Asp Ser Ser Pro Val Lys Ala Gly Val Glu Thr
170 175 180
Thr Thr Pro Ser Lys Gln Cys Asn Asn Lys Tyr Ala Ala Ser Ser
185 190 195
Tyr Leu Ser Leu Thr Pro Glu Gln Trp Lys Ser His Arg Ser Tyr
200 205 210
Ser Cys Gln Val Thr His Glu Gly Ser Thr Val Glu Lys Thr Val
215 220 225
Ala Pro Thr Glu Cys Ser
230

<210> 18
<211> 716
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3286805CD1

<400> 18
Met Asn Asn Phe Arg Ala Thr Ile Leu Phe Trp Ala Ala Ala Ala
1 5 10 15
Trp Ala Lys Ser Gly Lys Pro Ser Gly Glu Met Asp Glu Val Gly
20 25 30
Val Gln Lys Cys Lys Asn Ala Leu Lys Leu Pro Val Leu Glu Val
35 40 45
Leu Pro Gly Gly Gly Trp Asp Asn Leu Arg Asn Val Asp Met Gly
50 55 60
Arg Val Met Glu Leu Thr Tyr Ser Asn Cys Arg Thr Thr Glu Asp
65 70 75
Gly Gln Tyr Ile Ile Pro Asp Glu Ile Phe Thr Ile Pro Gln Lys
80 85 90
Gln Ser Asn Leu Glu Met Asn Ser Glu Ile Leu Glu Ser Trp Ala
95 100 105
Asn Tyr Gln Ser Ser Thr Ser Tyr Ser Ile Asn Thr Glu Leu Ser
110 115 120
Leu Phe Ser Lys Val Asn Gly Lys Phe Ser Thr Glu Phe Gln Arg

```
                125                 130                 135
Met Lys Thr Leu Gln Val Lys Asp Gln Ala Ile Thr Thr Arg Val
                140                 145                 150
Gln Val Arg Asn Leu Val Tyr Thr Val Lys Ile Asn Pro Thr Leu
                155                 160                 165
Glu Leu Ser Ser Gly Phe Arg Lys Glu Leu Leu Asp Ile Ser Asp
                170                 175                 180
Arg Leu Glu Asn Asn Gln Thr Arg Met Ala Thr Tyr Leu Ala Glu
                185                 190                 195
Leu Leu Val Leu Asn Tyr Gly Thr His Val Thr Thr Ser Val Asp
                200                 205                 210
Ala Gly Ala Ala Leu Ile Gln Glu Asp His Leu Arg Ala Ser Phe
                215                 220                 225
Leu Gln Asp Ser Gln Ser Ser Arg Ser Ala Val Thr Ala Ser Ala
                230                 235                 240
Gly Leu Ala Phe Gln Asn Thr Val Asn Phe Lys Phe Glu Glu Asn
                245                 250                 255
Tyr Thr Ser Gln Asn Val Leu Thr Lys Ser Tyr Leu Ser Asn Arg
                260                 265                 270
Thr Asn Ser Arg Val Gln Ser Ile Gly Gly Val Pro Phe Tyr Pro
                275                 280                 285
Gly Ile Thr Leu Gln Ala Trp Gln Gln Gly Ile Thr Asn His Leu
                290                 295                 300
Val Ala Ile Asp Arg Ser Gly Leu Pro Leu His Phe Phe Ile Asn
                305                 310                 315
Pro Asn Met Leu Pro Asp Leu Pro Gly Pro Leu Val Lys Lys Val
                320                 325                 330
Ser Lys Thr Val Glu Thr Ala Val Lys Arg Tyr Tyr Thr Phe Asn
                335                 340                 345
Thr Tyr Pro Gly Cys Thr Asp Leu Asn Ser Pro Asn Phe Asn Phe
                350                 355                 360
Gln Ala Asn Thr Asp Asp Gly Ser Cys Glu Gly Lys Met Thr Asn
                365                 370                 375
Phe Ser Phe Gly Gly Val Tyr Gln Glu Cys Thr Gln Leu Ser Gly
                380                 385                 390
Asn Arg Asp Val Leu Leu Cys Gln Lys Leu Glu Gln Lys Asn Pro
                395                 400                 405
Leu Thr Gly Asp Phe Ser Cys Pro Ser Gly Tyr Ser Pro Val His
                410                 415                 420
Leu Leu Ser Gln Ile His Glu Glu Gly Tyr Asn His Leu Glu Cys
                425                 430                 435
His Arg Lys Cys Thr Leu Leu Val Phe Cys Lys Thr Val Cys Glu
                440                 445                 450
Asp Val Phe Gln Val Ala Lys Ala Glu Phe Arg Ala Phe Trp Cys
                455                 460                 465
Val Ala Ser Ser Gln Val Pro Glu Asn Ser Gly Leu Leu Phe Gly
                470                 475                 480
Gly Leu Phe Ser Ser Lys Ser Ile Asn Pro Met Thr Asn Ala Gln
                485                 490                 495
Ser Cys Pro Ala Gly Tyr Phe Pro Leu Arg Leu Phe Glu Asn Leu
                500                 505                 510
Lys Val Cys Val Ser Gln Asp Tyr Glu Leu Gly Ser Arg Phe Ala
                515                 520                 525
Val Pro Phe Gly Gly Phe Phe Ser Cys Thr Val Gly Asn Pro Leu
                530                 535                 540
Val Asp Pro Ala Ile Ser Arg Asp Leu Gly Ala Pro Ser Leu Lys
                545                 550                 555
Lys Cys Pro Gly Gly Phe Ser Gln His Pro Ala Leu Ile Ser Asp
                560                 565                 570
Gly Cys Gln Val Ser Tyr Cys Val Lys Ser Gly Leu Phe Thr Gly
                575                 580                 585
Gly Ser Leu Pro Pro Ala Arg Leu Pro Pro Phe Thr Arg Pro Pro
                590                 595                 600
```

```
Leu Met Ser Gln Ala Ala Thr Asn Thr Val Ile Val Thr Asn Ser
                605                 610                 615
Glu Asn Ala Arg Ser Trp Ile Lys Asp Ser Gln Thr His Gln Trp
                620                 625                 630
Arg Leu Gly Glu Pro Ile Glu Leu Arg Arg Ala Met Asn Val Ile
                635                 640                 645
His Gly Asp Gly Gly Gly Leu Ser Gly Gly Ala Ala Ala Gly Val
                650                 655                 660
Thr Val Gly Val Thr Thr Ile Leu Ala Val Val Ile Thr Leu Ala
                665                 670                 675
Ile Tyr Gly Thr Arg Lys Phe Lys Lys Lys Ala Tyr Gln Ala Ile
                680                 685                 690
Glu Glu Arg Gln Ser Leu Val Pro Gly Thr Ala Ala Thr Gly Asp
                695                 700                 705
Thr Thr Tyr Gln Glu Gln Gly Gln Ser Pro Ala
                710                 715
```

<210> 19
<211> 519
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3506590CD1

<400> 19

```
Met Glu Phe Gly Leu Ser Trp Val Phe Leu Val Ala Leu Leu Arg
  1                 5                  10                  15
Gly Val Gln Cys Gln Val Gln Leu Val Glu Ser Gly Gly Gly Val
                 20                  25                  30
Val Gln Pro Gly Arg Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly
                 35                  40                  45
Phe Thr Phe Ser Ser Tyr Ala Met His Trp Val Arg Gln Ala Pro
                 50                  55                  60
Gly Lys Gly Leu Glu Trp Val Ala Val Ile Ser Tyr Asp Gly Ser
                 65                  70                  75
Asn Lys Tyr Tyr Ala Asp Ser Val Lys Gly Arg Phe Thr Ile Ser
                 80                  85                  90
Arg Asp Asn Ser Lys Asn Thr Leu Tyr Leu Gln Met Asn Ser Leu
                 95                 100                 105
Arg Ala Glu Asp Thr Ala Val Tyr Tyr Cys Ala Arg Ala Gly Glu
                110                 115                 120
Gly Ser Pro Asp Thr Leu Val Ala Phe Asp Ile Trp Gly Gln Gly
                125                 130                 135
Thr Met Val Thr Val Ser Ser Ala Ser Thr Lys Gly Pro Ser Val
                140                 145                 150
Phe Pro Leu Ala Pro Cys Ser Arg Ser Thr Ser Gly Gly Thr Ala
                155                 160                 165
Ala Leu Gly Cys Leu Val Lys Asp Tyr Phe Pro Glu Pro Val Thr
                170                 175                 180
Val Ser Trp Asn Ser Gly Ala Leu Thr Ser Gly Val His Thr Phe
                185                 190                 195
Pro Ala Val Leu Gln Ser Ser Gly Leu Tyr Ser Leu Ser Ser Val
                200                 205                 210
Val Thr Val Pro Ser Ser Ser Leu Gly Thr Gln Thr Tyr Thr Cys
                215                 220                 225
Asn Val Asn His Lys Pro Ser Asn Thr Lys Val Asp Lys Arg Val
                230                 235                 240
Glu Leu Lys Thr Pro Leu Gly Asp Thr Thr His Thr Cys Pro Arg
                245                 250                 255
Cys Pro Glu Pro Lys Ser Cys Asp Thr Pro Pro Pro Cys Pro Arg
                260                 265                 270
```

```
Cys Pro Glu Pro Lys Ser Cys Asp Thr Pro Pro Pro Cys Pro Arg
                275                 280                 285
Cys Pro Glu Pro Lys Ser Cys Asp Thr Pro Pro Pro Cys Pro Arg
                290                 295                 300
Cys Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser Val Phe Leu Phe
                305                 310                 315
Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg Thr Pro Glu
                320                 325                 330
Val Thr Cys Val Val Val Asp Val Ser His Glu Asp Pro Glu Val
                335                 340                 345
Gln Phe Lys Trp Tyr Val Asp Gly Val Glu Val His Asn Ala Lys
                350                 355                 360
Thr Lys Leu Arg Glu Glu Gln Tyr Asn Ser Thr Phe Arg Val Val
                365                 370                 375
Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys Glu
                380                 385                 390
Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile Glu
                395                 400                 405
Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val
                410                 415                 420
Tyr Thr Leu Pro Pro Ser Arg Glu Glu Met Thr Lys Asn Gln Val
                425                 430                 435
Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala
                440                 445                 450
Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Asn Thr
                455                 460                 465
Thr Pro Pro Met Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser
                470                 475                 480
Lys Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn Ile Phe
                485                 490                 495
Ser Cys Ser Val Met His Glu Ala Leu His Asn Arg Tyr Thr Gln
                500                 505                 510
Lys Ser Leu Ser Leu Ser Pro Gly Lys
                515
```

<210> 20
<211> 172
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 003600CD1

<400> 20

```
Met Leu Thr Glu Val Met Glu Val Trp His Gly Leu Val Ile Ala
  1               5                  10                  15
Val Val Ser Leu Phe Leu Gln Ala Cys Phe Leu Thr Ala Ile Asn
                20                  25                  30
Tyr Leu Leu Ser Arg His Met Ala His Lys Ser Glu Gln Ile Leu
                35                  40                  45
Lys Ala Ala Ser Leu Gln Val Pro Arg Pro Ser Pro Gly His His
                50                  55                  60
His Pro Pro Ala Val Lys Glu Met Lys Glu Thr Gln Thr Glu Arg
                65                  70                  75
Asp Ile Pro Met Ser Asp Ser Leu Tyr Arg His Asp Ser Asp Thr
                80                  85                  90
Pro Ser Asp Ser Leu Asp Ser Ser Cys Ser Ser Pro Pro Ala Cys
                95                 100                 105
Gln Ala Thr Glu Asp Val Asp Tyr Thr Gln Val Val Phe Ser Asp
                110                 115                 120
Pro Gly Glu Leu Lys Asn Asp Ser Pro Leu Asp Tyr Glu Asn Ile
                125                 130                 135
```

```
Lys Glu Ile Thr Asp Tyr Val Asn Val Asn Pro Glu Arg His Lys
                140                 145                 150
Pro Ser Phe Trp Tyr Phe Val Asn Pro Ala Leu Ser Glu Pro Ala
                155                 160                 165
Glu Tyr Asp Gln Val Ala Met
                170

<210> 21
<211> 314
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1251534CD1

<400> 21
Met Gly Leu Leu Asp Ser Glu Pro Gly Ser Val Leu Asn Val Val
  1                 5                  10                  15
Ser Thr Ala Leu Asn Asp Thr Val Glu Phe Tyr Arg Trp Thr Trp
                 20                  25                  30
Ser Ile Ala Asp Lys Arg Val Glu Asn Trp Pro Leu Met Gln Ser
                 35                  40                  45
Pro Trp Pro Thr Leu Ser Ile Ser Thr Leu Tyr Leu Leu Phe Val
                 50                  55                  60
Trp Leu Gly Pro Lys Trp Met Lys Asp Arg Glu Pro Phe Gln Met
                 65                  70                  75
Arg Leu Val Leu Ile Ile Tyr Asn Phe Gly Met Val Leu Leu Asn
                 80                  85                  90
Leu Phe Ile Phe Arg Glu Leu Phe Met Gly Ser Tyr Asn Ala Gly
                 95                 100                 105
Tyr Ser Tyr Ile Cys Gln Ser Val Asp Tyr Ser Asn Asn Val His
                110                 115                 120
Glu Val Arg Ile Ala Ala Ala Leu Trp Trp Tyr Phe Val Ser Lys
                125                 130                 135
Gly Val Glu Tyr Leu Asp Thr Val Phe Phe Ile Leu Arg Lys Lys
                140                 145                 150
Asn Asn Gln Val Ser Phe Leu His Val Tyr His His Cys Thr Met
                155                 160                 165
Phe Thr Leu Trp Trp Ile Gly Ile Lys Trp Val Ala Gly Gly Gln
                170                 175                 180
Ala Phe Phe Gly Ala Gln Leu Asn Ser Phe Ile His Val Ile Met
                185                 190                 195
Tyr Ser Tyr Tyr Gly Leu Thr Ala Phe Gly Pro Trp Ile Gln Lys
                200                 205                 210
Tyr Leu Trp Trp Lys Arg Tyr Leu Thr Met Leu Gln Leu Ile Gln
                215                 220                 225
Phe His Val Thr Ile Gly His Thr Ala Leu Ser Leu Tyr Thr Asp
                230                 235                 240
Cys Pro Phe Pro Lys Trp Met His Trp Ala Leu Ile Ala Tyr Ala
                245                 250                 255
Ile Ser Phe Ile Phe Leu Phe Leu Asn Phe Tyr Ile Arg Thr Tyr
                260                 265                 270
Lys Glu Pro Lys Lys Pro Lys Ala Gly Lys Thr Ala Met Asn Gly
                275                 280                 285
Ile Ser Ala Asn Gly Val Ser Lys Ser Glu Lys Gln Leu Met Ile
                290                 295                 300
Glu Asn Gly Lys Lys Gln Lys Asn Gly Lys Ala Lys Gly Asp
                305                 310

<210> 22
<211> 542
<212> PRT
```

<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1402211CD1

<400> 22

```
Met Asn Gly Lys Arg Pro Ala Glu Pro Gly Pro Ala Arg Val Gly
  1               5                  10                  15
Lys Lys Gly Lys Lys Glu Val Met Ala Glu Phe Ser Asp Ala Val
             20                  25                  30
Thr Glu Glu Thr Leu Lys Lys Gln Val Ala Glu Ala Trp Ser Arg
             35                  40                  45
Arg Thr Pro Phe Ser His Glu Val Ile Val Met Asp Met Asp Pro
             50                  55                  60
Phe Leu His Cys Val Ile Pro Asn Phe Ile Gln Ser Gln Asp Phe
             65                  70                  75
Leu Glu Gly Leu Gln Lys Glu Leu Met Asn Leu Asp Phe His Glu
             80                  85                  90
Lys Tyr Asn Asp Leu Tyr Lys Phe Gln Gln Ser Asp Asp Leu Lys
             95                 100                 105
Lys Arg Arg Glu Pro His Ile Ser Thr Leu Arg Lys Ile Leu Phe
            110                 115                 120
Glu Asp Phe Arg Ser Trp Leu Ser Asp Ile Ser Lys Ile Asp Leu
            125                 130                 135
Glu Ser Thr Ile Asp Met Ser Cys Ala Lys Tyr Glu Phe Thr Asp
            140                 145                 150
Ala Leu Leu Cys His Asp Asp Glu Leu Glu Gly Arg Arg Ile Ala
            155                 160                 165
Phe Ile Leu Tyr Leu Val Pro Pro Trp Asp Arg Ser Met Gly Gly
            170                 175                 180
Thr Leu Asp Leu Tyr Ser Ile Asp Glu His Phe Gln Pro Lys Gln
            185                 190                 195
Ile Val Lys Ser Leu Ile Pro Ser Trp Asn Lys Leu Val Phe Phe
            200                 205                 210
Glu Val Ser Pro Val Ser Phe His Gln Val Ser Glu Val Leu Ser
            215                 220                 225
Glu Glu Lys Ser Arg Leu Ser Ile Ser Gly Trp Phe His Gly Pro
            230                 235                 240
Ser Leu Thr Arg Pro Pro Asn Tyr Phe Glu Pro Pro Ile Pro Arg
            245                 250                 255
Ser Pro His Ile Pro Gln Asp His Glu Ile Leu Tyr Asp Trp Ile
            260                 265                 270
Asn Pro Thr Tyr Leu Asp Met Asp Tyr Gln Val Gln Ile Gln Glu
            275                 280                 285
Glu Phe Glu Glu Ser Ser Glu Ile Leu Leu Lys Glu Phe Leu Lys
            290                 295                 300
Pro Glu Lys Phe Thr Lys Val Cys Glu Ala Leu Glu His Gly His
            305                 310                 315
Val Glu Trp Ser Ser Arg Gly Pro Pro Asn Lys Arg Phe Tyr Glu
            320                 325                 330
Lys Ala Glu Glu Ser Lys Leu Pro Glu Ile Leu Lys Glu Cys Met
            335                 340                 345
Lys Leu Phe Arg Ser Glu Ala Leu Phe Leu Leu Leu Ser Asn Phe
            350                 355                 360
Thr Gly Leu Lys Leu His Phe Leu Ala Pro Ser Glu Glu Asp Glu
            365                 370                 375
Met Asn Asp Lys Lys Glu Ala Glu Thr Thr Asp Ile Thr Glu Glu
            380                 385                 390
Gly Thr Ser His Ser Pro Pro Glu Pro Glu Asn Asn Gln Met Ala
            395                 400                 405
Ile Ser Asn Asn Ser Gln Gln Ser Asn Glu Gln Thr Asp Pro Glu
            410                 415                 420
```

```
Pro Glu Glu Asn Glu Thr Lys Lys Glu Ser Ser Val Pro Met Cys
                425                 430                 435
Gln Gly Glu Leu Arg His Trp Lys Thr Gly His Tyr Thr Leu Ile
                440                 445                 450
His Asp His Ser Lys Ala Glu Phe Ala Leu Asp Leu Ile Leu Tyr
                455                 460                 465
Cys Gly Cys Glu Gly Trp Glu Pro Glu Tyr Gly Gly Phe Thr Ser
                470                 475                 480
Tyr Ile Ala Lys Gly Glu Asp Glu Glu Leu Leu Thr Val Asn Pro
                485                 490                 495
Glu Ser Asn Ser Leu Ala Leu Val Tyr Arg Asp Arg Glu Thr Leu
                500                 505                 510
Lys Phe Val Lys His Ile Asn His Arg Ser Leu Glu Gln Lys Lys
                515                 520                 525
Thr Phe Pro Asn Arg Thr Gly Phe Trp Asp Phe Ser Phe Ile Tyr
                530                 535                 540
Tyr Glu
```

<210> 23
<211> 715
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1623474CD1

<400> 23

```
Met Pro Ala Glu Ser Gly Lys Arg Phe Lys Pro Ser Lys Tyr Val
  1               5                  10                  15
Pro Val Ser Ala Ala Ala Ile Phe Leu Val Gly Ala Thr Thr Leu
             20                  25                  30
Phe Phe Ala Phe Thr Cys Pro Gly Leu Ser Leu Tyr Val Ser Pro
             35                  40                  45
Ala Val Pro Ile Tyr Asn Ala Ile Met Phe Leu Phe Val Leu Ala
             50                  55                  60
Asn Phe Ser Met Ala Thr Phe Met Asp Pro Gly Ile Phe Pro Arg
             65                  70                  75
Ala Glu Glu Asp Glu Asp Lys Glu Asp Asp Phe Arg Ala Pro Leu
             80                  85                  90
Tyr Lys Thr Val Glu Ile Lys Gly Ile Gln Val Arg Met Lys Trp
             95                 100                 105
Cys Ala Thr Cys Arg Phe Tyr Arg Pro Pro Arg Cys Ser His Cys
            110                 115                 120
Ser Val Cys Asp Asn Cys Val Glu Glu Phe Asp His His Cys Pro
            125                 130                 135
Trp Val Asn Asn Cys Ile Gly Arg Arg Asn Tyr Arg Tyr Phe Phe
            140                 145                 150
Leu Phe Leu Leu Ser Leu Thr Ala His Ile Met Gly Val Phe Gly
            155                 160                 165
Phe Gly Leu Leu Tyr Val Leu Tyr His Ile Glu Glu Leu Ser Gly
            170                 175                 180
Val Arg Thr Ala Val Thr Met Ala Val Met Cys Val Ala Gly Leu
            185                 190                 195
Phe Phe Ile Pro Val Ala Gly Leu Thr Gly Phe His Val Val Leu
            200                 205                 210
Val Ala Arg Gly Arg Thr Thr Asn Glu Gln Val Thr Gly Lys Phe
            215                 220                 225
Arg Gly Gly Val Asn Pro Phe Thr Asn Gly Cys Cys Asn Asn Val
            230                 235                 240
Ser Arg Val Leu Cys Ser Ser Pro Ala Pro Arg Tyr Leu Gly Arg
            245                 250                 255
```

```
Pro Lys Lys Glu Lys Thr Ile Val Ile Arg Pro Pro Phe Leu Arg
                260                 265                 270
Pro Glu Val Ser Asp Gly Gln Ile Thr Val Lys Ile Met Asp Asn
                275                 280                 285
Gly Ile Gln Gly Glu Leu Arg Arg Thr Lys Ser Lys Gly Ser Leu
                290                 295                 300
Glu Ile Thr Glu Ser Gln Ser Ala Asp Ala Glu Pro Pro Pro Pro
                305                 310                 315
Pro Lys Pro Asp Leu Ser Arg Tyr Thr Gly Leu Arg Thr His Leu
                320                 325                 330
Gly Leu Ala Thr Asn Glu Asp Ser Ser Leu Leu Ala Lys Asp Ser
                335                 340                 345
Pro Pro Thr Pro Thr Met Tyr Lys Tyr Arg Pro Gly Tyr Ser Ser
                350                 355                 360
Ser Ser Thr Ser Ala Ala Met Pro His Ser Ser Ser Ala Lys Leu
                365                 370                 375
Ser Arg Gly Asp Ser Leu Lys Glu Pro Thr Ser Ile Ala Glu Ser
                380                 385                 390
Ser Arg His Pro Ser Tyr Arg Ser Glu Pro Ser Leu Glu Pro Glu
                395                 400                 405
Ser Phe Arg Ser Pro Thr Phe Gly Lys Ser Phe His Phe Asp Pro
                410                 415                 420
Leu Ser Ser Gly Ser Arg Ser Ser Ser Leu Lys Ser Ala Gln Gly
                425                 430                 435
Thr Gly Phe Glu Leu Gly Gln Leu Gln Ser Ile Arg Ser Glu Gly
                440                 445                 450
Thr Thr Ser Thr Ser Tyr Lys Ser Leu Ala Asn Gln Thr Arg Asn
                455                 460                 465
Gly Ser Leu Ser Tyr Asp Ser Leu Leu Thr Pro Ser Asp Ser Pro
                470                 475                 480
Asp Phe Glu Ser Val Gln Ala Gly Pro Glu Pro Asp Pro Pro Leu
                485                 490                 495
Gly Tyr Thr Ser Pro Phe Leu Ser Ala Arg Leu Ala Gln Gln Arg
                500                 505                 510
Glu Ala Glu Arg His Pro Arg Leu Val Pro Thr Gly Pro Thr His
                515                 520                 525
Arg Glu Pro Ser Pro Val Arg Tyr Asp Asn Leu Ser Arg His Ile
                530                 535                 540
Val Ala Ser Leu Gln Glu Arg Glu Lys Leu Leu Arg Gln Ser Pro
                545                 550                 555
Pro Leu Pro Gly Arg Glu Glu Glu Pro Gly Leu Gly Asp Ser Gly
                560                 565                 570
Ile Gln Ser Thr Pro Gly Ser Gly His Ala Pro Arg Thr Ser Ser
                575                 580                 585
Ser Ser Asp Asp Ser Lys Arg Ser Pro Leu Gly Lys Thr Pro Leu
                590                 595                 600
Gly Arg Pro Ala Val Pro Arg Phe Gly Lys Pro Asp Gly Leu Arg
                605                 610                 615
Gly Arg Gly Val Gly Ser Pro Glu Pro Gly Pro Thr Ala Pro Tyr
                620                 625                 630
Leu Gly Arg Ser Met Ser Tyr Ser Ser Gln Lys Ala Gln Pro Gly
                635                 640                 645
Val Ser Glu Thr Glu Glu Val Ala Leu Gln Pro Leu Leu Thr Pro
                650                 655                 660
Lys Asp Glu Val Gln Leu Lys Thr Thr Tyr Ser Lys Ser Asn Gly
                665                 670                 675
Gln Pro Lys Ser Leu Gly Ser Ala Ser Pro Gly Pro Gly Gln Pro
                680                 685                 690
Pro Leu Ser Ser Pro Thr Arg Gly Gly Val Lys Lys Val Ser Gly
                695                 700                 705
Val Gly Gly Thr Thr Tyr Glu Ile Ser Val
                710                 715
```

<210> 24
<211> 469
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1706443CD1

<400> 24
Met Gly Arg Val Arg Arg Ile Tyr Pro Gln Leu Leu Leu Ala Leu
1 5 10 15
Leu Ile Gln Val His Tyr His Ile Gly Leu Asn Leu Pro Gly Cys
20 25 30
Val Ala Pro Pro Lys Asp Thr Lys Lys Gly Ala Gln Pro Ser Pro
35 40 45
Phe Val Pro Val Arg Trp Val Val Lys Val Val Lys Thr Leu Leu
50 55 60
Leu Arg Met Gly Cys Ser Tyr Glu Thr Thr Phe Leu Glu Asp Gln
65 70 75
Gly Gly Trp Glu Leu Met Glu Gln Val Glu Ser His His Arg Gly
80 85 90
Val Ala Leu Leu Ala Arg Ala Met Val Gln Tyr Ser Cys Gln Glu
95 100 105
Leu Cys Arg Ile Leu Tyr Leu Leu Ile Pro Leu Leu Glu Arg Gly
110 115 120
Asp Glu Lys His Arg Ile Thr Ala Thr Ala Phe Phe Val Glu Leu
125 130 135
Leu Gln Met Glu Gln Val Arg Arg Ile Pro Glu Glu Tyr Ser Leu
140 145 150
Gly Arg Met Ala Glu Gly Leu Ser His His Asp Pro Ile Met Lys
155 160 165
Val Leu Ser Ile Arg Gly Leu Val Ile Leu Ala Arg Arg Ser Glu
170 175 180
Lys Thr Ala Lys Val Lys Ala Leu Leu Pro Ser Met Val Lys Gly
185 190 195
Leu Lys Asn Met Asp Gly Met Leu Val Val Glu Ala Val His Asn
200 205 210
Leu Lys Ala Val Phe Lys Gly Arg Asp Gln Lys Leu Met Asp Ser
215 220 225
Ala Val Tyr Val Glu Met Leu Gln Ile Leu Leu Pro His Phe Ser
230 235 240
Asp Ala Arg Glu Asp Val Arg Ser Ser Cys Ile Asn Leu Tyr Gly
245 250 255
Lys Val Val Gln Lys Leu Arg Ala Pro Arg Thr Gln Ala Met Glu
260 265 270
Glu Gln Leu Val Ser Thr Leu Val Pro Leu Leu Leu Thr Met Gln
275 280 285
Glu Gly Asn Ser Lys Val Ser Gln Lys Cys Val Lys Thr Leu Leu
290 295 300
Arg Cys Ser Tyr Phe Met Ala Trp Glu Leu Pro Lys Arg Ala Tyr
305 310 315
Ser Arg Lys Pro Trp Asp Asn Gln Gln Gln Thr Val Ala Lys Ile
320 325 330
Cys Lys Cys Leu Val Asn Thr His Arg Asp Ser Ala Phe Ile Phe
335 340 345
Leu Ser Gln Ser Leu Glu Tyr Ala Lys Asn Ser Arg Ala Ser Leu
350 355 360
Arg Lys Cys Ser Val Met Phe Ile Gly Ser Leu Val Pro Cys Met
365 370 375
Glu Ser Ile Met Thr Glu Asp Arg Leu Asn Glu Val Lys Ala Ala
380 385 390
Leu Asp Asn Leu Arg His Asp Pro Glu Ala Ser Val Cys Ile Tyr

```
                395                 400                 405
Ala Ala Gln Val Gln Asp His Ile Leu Ala Ser Cys Trp Gln Asn
                410                 415                 420
Ser Trp Leu Pro His Gly Asn Ser Trp Val Cys Tyr Ser Ala Thr
                425                 430                 435
Thr His Arg Trp Ser Pro Ser Cys Glu Asn Leu Pro Thr Ser His
                440                 445                 450
Gln Arg Arg Ser Trp Ile Met Gln Ala Leu Gly Ser Trp Lys Met
                455                 460                 465
Ser Leu Lys Lys
```

<210> 25
<211> 274
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1748627CD1

<400> 25

```
Met Pro Arg Ala Glu Pro Arg Ala Thr Leu Gly Glu Gln Glu Lys
  1               5                  10                  15
Ala Gly Leu Pro Leu Gly Ala Trp Arg Leu Tyr Leu Leu Arg His
                20                  25                  30
Phe Arg Lys Gln Thr Glu Leu Arg Arg Ser Gly Ser Arg Asp Val
                35                  40                  45
Thr Gly Ala Leu Leu Val Ala Ala Ala Val Ala Ser Glu Ala Val
                50                  55                  60
Gly Ser Leu Arg Val Ala Glu Gly Gly Pro Asn Thr Leu Leu Leu
                65                  70                  75
Gln Val Leu Arg Ser Trp Pro Trp Cys Asn Lys Glu Leu Lys Thr
                80                  85                  90
Met Glu Glu Arg Lys Val Lys Arg Arg Ser Pro Lys Ser Phe Ser
                95                 100                 105
Ala His Cys Thr Gln Val Val Asn Ala Lys Lys Asn Ala Ile Pro
               110                 115                 120
Val Ser Lys Ser Thr Gly Phe Ser Asn Pro Ala Ser Gln Ser Thr
               125                 130                 135
Ser Gln Arg Pro Lys Leu Lys Arg Val Met Lys Glu Lys Thr Lys
               140                 145                 150
Pro Gln Gly Gly Glu Gly Lys Gly Ala Gln Ser Thr Pro Ile Gln
               155                 160                 165
His Ser Phe Leu Thr Asp Val Ser Asp Val Gln Glu Met Glu Arg
               170                 175                 180
Gly Leu Leu Ser Leu Leu Asn Asp Phe His Ser Gly Lys Leu Gln
               185                 190                 195
Ala Phe Gly Asn Glu Cys Ser Ile Glu Gln Met Glu His Val Arg
               200                 205                 210
Gly Met Gln Glu Lys Leu Ala Arg Leu Asn Leu Glu Leu Tyr Gly
               215                 220                 225
Glu Leu Glu Glu Leu Pro Glu Asp Lys Arg Lys Thr Ala Ser Asp
               230                 235                 240
Ser Asn Leu Asp Arg Leu Leu Ser Asp Leu Glu Glu Leu Asn Ser
               245                 250                 255
Ser Ile Gln Lys Leu His Leu Ala Asp Ala Gln Asp Val Pro Asn
               260                 265                 270
Thr Ser Ala Ser
```

<210> 26
<211> 154

<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1818332CD1

<400> 26
Met Ala Gly Pro Val Lys Asp Arg Glu Ala Phe Gln Arg Leu Asn
1 5 10 15
Phe Leu Tyr Gln Ala Ala His Cys Val Leu Ala Gln Asp Pro Glu
20 25 30
Asn Gln Ala Leu Ala Arg Phe Tyr Cys Tyr Thr Glu Arg Thr Ile
35 40 45
Ala Lys Arg Leu Val Leu Arg Arg Asp Pro Ser Val Lys Arg Thr
50 55 60
Leu Cys Arg Gly Cys Ser Ser Leu Leu Val Pro Gly Leu Thr Cys
65 70 75
Thr Gln Arg Gln Arg Arg Cys Arg Gly Gln Arg Trp Thr Val Gln
80 85 90
Thr Cys Leu Thr Cys Gln Arg Ser Gln Arg Phe Leu Asn Asp Pro
95 100 105
Gly His Leu Leu Trp Gly Asp Arg Pro Glu Ala Gln Leu Gly Ser
110 115 120
Gln Ala Asp Ser Lys Pro Leu Gln Pro Leu Pro Asn Thr Ala His
125 130 135
Ser Ile Ser Asp Arg Leu Pro Glu Glu Lys Met Gln Thr Gln Gly
140 145 150
Ser Ser Asn Gln

<210> 27
<211> 102
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1822832CD1

<400> 27
Met Lys Phe Asp Trp Val Met Gly Leu Arg Ser Ile Thr Leu Lys
1 5 10 15
Asn Ser Ser Thr Gly Arg Gly Asp Gly Pro Lys Gln His Leu Gln
20 25 30
Ala Asp Pro Met Leu Ile Ile Arg Ala Arg Thr Leu Ser Leu Ser
35 40 45
Val Ser Leu Ser Val Ser Pro Leu Gly Leu Thr Pro His Trp Thr
50 55 60
Pro Leu His Pro Cys Pro Ser His Asn Thr Ala Ala Val Ser Ser
65 70 75
Ala Cys Leu Trp Glu Ser Pro Leu Phe Ser Ser Val Phe Phe Ser
80 85 90
Ser Cys Pro Ile Thr Pro Cys Thr Ser Pro Phe Pro
95 100

<210> 28
<211> 113
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature

<223> Incyte ID No: 1832219CD1

<400> 28
Met Ala Gly Pro Ala Ala Ala Phe Arg Arg Leu Gly Ala Leu Ser
1 5 10 15
Gly Ala Ala Ala Leu Gly Phe Ala Ser Tyr Gly Ala His Gly Ala
20 25 30
Gln Phe Pro Asp Ala Tyr Gly Lys Glu Leu Phe Asp Lys Ala Asn
35 40 45
Lys His His Phe Leu His Ser Leu Ala Leu Leu Gly Val Pro His
50 55 60
Cys Arg Lys Pro Leu Trp Ala Gly Leu Leu Leu Ala Ser Gly Thr
65 70 75
Thr Leu Phe Cys Thr Ser Phe Tyr Tyr Gln Ala Leu Ser Gly Asp
80 85 90
Pro Ser Ile Gln Thr Leu Ala Pro Ala Gly Gly Thr Leu Leu Leu
95 100 105
Leu Gly Trp Leu Ala Leu Ala Leu
110

<210> 29
<211> 313
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1899010CD1

<400> 29
Met Ala Leu Leu Val Asp Arg Val Arg Gly His Trp Arg Ile Ala
1 5 10 15
Ala Gly Leu Leu Phe Asn Leu Leu Val Ser Ile Cys Ile Val Phe
20 25 30
Leu Asn Lys Trp Ile Tyr Val Tyr His Gly Phe Pro Asn Met Ser
35 40 45
Leu Thr Leu Val His Phe Val Val Thr Trp Leu Gly Leu Tyr Ile
50 55 60
Cys Gln Lys Leu Asp Ile Phe Ala Pro Lys Ser Leu Pro Pro Ser
65 70 75
Arg Leu Leu Leu Leu Ala Leu Ser Phe Cys Gly Phe Val Val Phe
80 85 90
Thr Asn Leu Ser Leu Gln Asn Asn Thr Ile Gly Thr Tyr Gln Leu
95 100 105
Ala Lys Ala Met Thr Thr Pro Val Ile Ile Ala Ile Gln Thr Phe
110 115 120
Cys Tyr Gln Lys Thr Phe Ser Thr Arg Ile Gln Leu Thr Leu Ile
125 130 135
Pro Ile Thr Leu Gly Val Ile Leu Asn Ser Tyr Tyr Asp Val Lys
140 145 150
Phe Asn Phe Leu Gly Met Val Phe Ala Ala Leu Gly Val Leu Val
155 160 165
Thr Ser Leu Tyr Gln Val Trp Val Gly Ala Lys Gln His Glu Leu
170 175 180
Gln Val Asn Ser Met Gln Leu Leu Tyr Tyr Gln Ala Pro Met Ser
185 190 195
Ser Ala Met Leu Leu Val Ala Val Pro Phe Phe Glu Pro Val Phe
200 205 210
Gly Glu Gly Gly Ile Phe Gly Pro Trp Ser Val Ser Ala Leu Leu
215 220 225
Met Val Leu Leu Ser Gly Val Ile Ala Phe Met Val Asn Leu Ser
230 235 240
Ile Tyr Trp Ile Ile Gly Asn Thr Ser Pro Val Thr Tyr Asn Met

```
                245                 250                 255
Phe Gly His Phe Lys Phe Cys Ile Thr Leu Phe Gly Gly Tyr Val
                260                 265                 270
Leu Phe Lys Asp Pro Leu Ser Ile Asn Gln Ala Leu Gly Ile Leu
                275                 280                 285
Cys Thr Leu Phe Gly Ile Leu Ala Tyr Thr His Phe Lys Leu Ser
                290                 295                 300
Glu Gln Glu Gly Ser Arg Ser Lys Leu Ala Gln Arg Pro
                305                 310
```

<210> 30
<211> 195
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2008768CD1

<400> 30

```
Met Ala Pro Lys Ala Ala Lys Gly Ala Lys Pro Glu Pro Ala Pro
  1                 5                  10                  15
Ala Pro Pro Pro Pro Gly Ala Lys Pro Glu Glu Asp Lys Lys Asp
                 20                  25                  30
Gly Lys Glu Pro Ser Asp Lys Pro Gln Lys Ala Val Gln Asp His
                 35                  40                  45
Lys Glu Pro Ser Asp Lys Pro Gln Lys Ala Val Gln Pro Lys His
                 50                  55                  60
Glu Val Gly Thr Arg Arg Gly Cys Arg Arg Tyr Arg Trp Glu Leu
                 65                  70                  75
Lys Asp Ser Asn Lys Glu Phe Trp Leu Leu Gly His Ala Glu Ile
                 80                  85                  90
Lys Ile Arg Ser Leu Asp Leu Phe Asn Asp Leu Ile Ala Cys Ala
                 95                 100                 105
Phe Leu Val Gly Ala Val Val Phe Ala Val Arg Ser Arg Arg Ser
                110                 115                 120
Met Asn Leu His Tyr Leu Leu Ala Val Ile Leu Ile Gly Ala Ala
                125                 130                 135
Gly Val Phe Ala Phe Ile Asp Val Cys Leu Gln Arg Asn His Phe
                140                 145                 150
Arg Gly Lys Lys Ala Lys Lys His Met Leu Val Pro Pro Pro Gly
                155                 160                 165
Lys Glu Lys Gly Pro Gln Gln Gly Lys Gly Pro Glu Pro Ala Lys
                170                 175                 180
Pro Pro Glu Pro Gly Lys Pro Pro Gly Pro Ala Lys Gly Lys Lys
                185                 190                 195
```

<210> 31
<211> 350
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2070984CD1

<400> 31

```
Met Asn Leu Arg Gly Leu Phe Gln Asp Phe Asn Pro Ser Lys Phe
  1                 5                  10                  15
Leu Ile Tyr Ala Cys Leu Leu Leu Phe Ser Val Leu Leu Ala Leu
                 20                  25                  30
Arg Leu Asp Gly Ile Ile Gln Trp Ser Tyr Trp Ala Val Phe Ala
```

```
                35                  40                  45
Pro Ile Trp Leu Trp Lys Leu Met Val Ile Val Gly Ala Ser Val
                50                  55                  60
Gly Thr Gly Val Trp Ala Arg Asn Pro Gln Tyr Arg Ala Glu Gly
                65                  70                  75
Glu Thr Cys Val Glu Phe Lys Ala Met Leu Ile Ala Val Gly Ile
                80                  85                  90
His Leu Leu Leu Leu Met Phe Glu Val Leu Val Cys Asp Arg Ile
                95                  100                 105
Glu Arg Gly Ser His Phe Trp Leu Leu Val Phe Met Pro Leu Phe
                110                 115                 120
Phe Val Ser Pro Val Ser Val Ala Ala Cys Val Trp Gly Phe Arg
                125                 130                 135
His Asp Arg Ser Leu Glu Leu Glu Ile Leu Cys Ser Val Asn Ile
                140                 145                 150
Leu Gln Phe Ile Phe Ile Ala Leu Arg Leu Asp Lys Ile Ile His
                155                 160                 165
Trp Pro Trp Leu Val Val Cys Val Pro Leu Trp Ile Leu Met Ser
                170                 175                 180
Phe Leu Cys Leu Val Val Leu Tyr Tyr Ile Val Trp Ser Val Leu
                185                 190                 195
Phe Leu Arg Ser Met Asp Val Ile Ala Glu Gln Arg Arg Thr His
                200                 205                 210
Ile Thr Met Ala Leu Ser Trp Met Thr Ile Val Val Pro Leu Leu
                215                 220                 225
Thr Phe Glu Ile Leu Leu Val His Lys Leu Asp Gly His Asn Ala
                230                 235                 240
Phe Ser Cys Ile Pro Ile Phe Val Pro Leu Trp Leu Ser Leu Ile
                245                 250                 255
Thr Leu Met Ala Thr Thr Phe Gly Gln Lys Gly Gly Asn His Trp
                260                 265                 270
Trp Phe Gly Ile Arg Lys Asp Phe Cys Gln Phe Leu Leu Glu Ile
                275                 280                 285
Phe Pro Phe Leu Arg Glu Tyr Gly Asn Ile Ser Tyr Asp Leu His
                290                 295                 300
His Glu Asp Asn Glu Glu Thr Glu Glu Thr Pro Val Pro Glu Pro
                305                 310                 315
Pro Lys Ile Ala Pro Met Phe Arg Lys Lys Ala Arg Val Val Ile
                320                 325                 330
Thr Gln Ser Pro Gly Lys Tyr Val Leu Pro Pro Pro Lys Leu Asn
                335                 340                 345
Ile Glu Met Pro Asp
                350
```

<210> 32
<211> 360
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2193240CD1

<400> 32

```
Met Ser Leu Leu Ala Val Ser Arg Arg Ala Gln Lys His Ala Leu
  1                 5                  10                  15
Lys Ala Asn Leu Ile Asp Asn Cys Met Glu Gln Met Lys His Ile
                20                  25                  30
Asn Ala Gln Leu Asn Leu Asp Ser Leu Arg Pro Gly Lys Ala Ala
                35                  40                  45
Leu Lys Lys Lys Glu Asp Gly Val Ile Lys Glu Leu Ser Ile Ala
                50                  55                  60
Met Gln Leu Leu Arg Asn Cys Leu Tyr Gln Asn Glu Glu Cys Lys
```

```
                 65                  70                  75
Glu Ala Ala Leu Glu Ala His Leu Val Pro Val Leu His Ser Leu
                 80                  85                  90
Trp Pro Trp Ile Leu Met Asp Asp Ser Leu Met Gln Ile Ser Leu
                 95                 100                 105
Gln Leu Leu Cys Val Tyr Thr Ala Asn Phe Pro Asn Gly Cys Ser
                110                 115                 120
Ser Leu Cys Trp Ser Ser Cys Gly Gln His Pro Val Gln Ala Thr
                125                 130                 135
His Arg Gly Ala Val Ser Asn Ser Leu Met Leu Cys Ile Leu Lys
                140                 145                 150
Leu Ala Ser Gln Met Pro Leu Glu Asn Thr Thr Val Gln Gln Met
                155                 160                 165
Val Phe Met Leu Leu Ser Asn Leu Ala Leu Ser His Asp Cys Lys
                170                 175                 180
Gly Val Ile Gln Lys Ser Asn Phe Leu Gln Asn Phe Leu Ser Leu
                185                 190                 195
Ala Leu Pro Lys Gly Gly Asn Lys His Leu Ser Asn Leu Thr Ile
                200                 205                 210
Leu Trp Leu Lys Leu Leu Leu Asn Ile Ser Ser Gly Glu Asp Gly
                215                 220                 225
Gln Gln Met Ile Leu Arg Leu Asp Gly Cys Leu Asp Leu Leu Thr
                230                 235                 240
Glu Met Ser Lys Tyr Lys His Lys Ser Ser Pro Leu Leu Pro Leu
                245                 250                 255
Leu Ile Phe His Asn Val Cys Phe Ser Pro Ala Asn Lys Pro Lys
                260                 265                 270
Ile Leu Ala Asn Glu Lys Val Ile Thr Val Leu Ala Ala Cys Leu
                275                 280                 285
Glu Ser Glu Asn Gln Asn Ala Gln Arg Ile Gly Ala Ala Ala Leu
                290                 295                 300
Trp Ala Leu Ile Tyr Asn Tyr Gln Lys Ala Lys Thr Ala Leu Lys
                305                 310                 315
Ser Pro Ser Val Lys Arg Arg Val Asp Glu Ala Tyr Ser Leu Ala
                320                 325                 330
Lys Lys Thr Phe Pro Asn Ser Glu Ala Asn Pro Leu Asn Ala Tyr
                335                 340                 345
Tyr Leu Lys Cys Leu Glu Asn Leu Val Gln Leu Leu Asn Ser Ser
                350                 355                 360
```

<210> 33
<211> 559
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2235177CD1

<400> 33

```
Met Gly Ser Arg Ile Lys Gln Asn Pro Glu Thr Thr Phe Glu Val
  1                 5                  10                  15
Tyr Val Glu Val Ala Tyr Pro Arg Thr Gly Gly Thr Leu Ser Asp
                 20                  25                  30
Pro Glu Val Gln Arg Gln Phe Pro Glu Asp Tyr Ser Asp Gln Glu
                 35                  40                  45
Val Leu Gln Thr Leu Thr Lys Phe Cys Phe Pro Phe Tyr Val Asp
                 50                  55                  60
Ser Leu Thr Val Ser Gln Val Gly Gln Asn Phe Thr Phe Val Leu
                 65                  70                  75
Thr Asp Ile Asp Ser Lys Gln Arg Phe Gly Phe Cys Arg Leu Ser
                 80                  85                  90
```

```
Ser Gly Ala Lys Ser Cys Phe Cys Ile Leu Ser Tyr Leu Pro Trp
                 95                  100                 105
Phe Glu Val Phe Tyr Lys Leu Leu Asn Ile Leu Ala Asp Tyr Thr
                110                 115                 120
Thr Lys Arg Gln Glu Asn Gln Trp Asn Glu Leu Leu Glu Thr Leu
                125                 130                 135
His Lys Leu Pro Ile Pro Asp Pro Gly Val Ser Val His Leu Ser
                140                 145                 150
Val His Ser Tyr Phe Thr Val Pro Asp Thr Arg Glu Leu Pro Ser
                155                 160                 165
Ile Pro Glu Asn Arg Asn Leu Thr Glu Tyr Phe Val Ala Val Asp
                170                 175                 180
Val Asn Asn Met Leu His Leu Tyr Ala Ser Met Leu Tyr Glu Arg
                185                 190                 195
Arg Ile Leu Ile Ile Cys Ser Lys Leu Ser Thr Leu Thr Ala Cys
                200                 205                 210
Ile His Gly Ser Ala Ala Met Leu Tyr Pro Met Tyr Trp Gln His
                215                 220                 225
Val Tyr Ile Pro Val Leu Pro Pro His Leu Leu Asp Tyr Cys Cys
                230                 235                 240
Ala Pro Met Pro Tyr Leu Ile Gly Ile His Leu Ser Leu Met Glu
                245                 250                 255
Lys Val Arg Asn Met Ala Leu Asp Asp Val Val Ile Leu Asn Val
                260                 265                 270
Asp Thr Asn Thr Leu Glu Thr Pro Phe Asp Asp Leu Gln Ser Leu
                275                 280                 285
Pro Asn Asp Val Ile Ser Ser Leu Lys Asn Arg Leu Lys Lys Val
                290                 295                 300
Ser Thr Thr Thr Gly Asp Gly Val Ala Arg Ala Phe Leu Lys Ala
                305                 310                 315
Gln Ala Ala Phe Phe Gly Ser Tyr Arg Asn Ala Leu Lys Ile Glu
                320                 325                 330
Pro Glu Glu Pro Ile Thr Phe Cys Glu Glu Ala Phe Val Ser His
                335                 340                 345
Tyr Arg Ser Gly Ala Met Arg Gln Phe Leu Gln Asn Ala Thr Gln
                350                 355                 360
Leu Gln Leu Phe Lys Gln Phe Ile Asp Gly Arg Leu Asp Leu Leu
                365                 370                 375
Asn Ser Gly Glu Gly Phe Ser Asp Val Phe Glu Glu Glu Ile Asn
                380                 385                 390
Met Gly Glu Tyr Ala Gly Ser Asp Lys Leu Tyr His Gln Trp Leu
                395                 400                 405
Ser Thr Val Arg Lys Gly Ser Gly Ala Ile Leu Asn Thr Val Lys
                410                 415                 420
Thr Lys Ala Asn Pro Ala Met Lys Thr Val Tyr Lys Phe Ala Lys
                425                 430                 435
Asp His Ala Lys Met Gly Ile Lys Glu Val Lys Asn Arg Leu Lys
                440                 445                 450
Gln Lys Asp Ile Ala Glu Asn Gly Cys Ala Pro Thr Pro Glu Glu
                455                 460                 465
Gln Leu Pro Lys Thr Ala Pro Ser Pro Leu Val Glu Ala Lys Asp
                470                 475                 480
Pro Lys Leu Arg Glu Asp Arg Arg Pro Ile Thr Val His Phe Gly
                485                 490                 495
Gln Val Arg Pro Pro Arg Pro His Val Val Lys Arg Pro Lys Ser
                500                 505                 510
Asn Ile Ala Val Glu Gly Arg Arg Thr Ser Val Pro Ser Pro Glu
                515                 520                 525
Gln Asn Thr Ile Ala Thr Pro Ala Thr Leu His Ile Leu Gln Lys
                530                 535                 540
Ser Ile Thr His Phe Ala Ala Lys Phe Pro Thr Arg Gly Trp Thr
                545                 550                 555
Ser Ser Ser His
```

<210> 34
<211> 198
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2416227CD1

<400> 34
Met Ala Leu Arg His Leu Ala Leu Leu Ala Gly Leu Leu Val Gly
1 5 10 15
Val Ala Ser Lys Ser Met Glu Asn Thr Ala Gln Leu Pro Glu Cys
20 25 30
Cys Val Asp Val Val Gly Val Asn Ala Ser Cys Pro Gly Ala Ser
35 40 45
Leu Cys Gly Pro Gly Cys Tyr Arg Arg Trp Asn Ala Asp Gly Ser
50 55 60
Ala Ser Cys Val Arg Cys Gly Asn Gly Thr Leu Pro Ala Tyr Asn
65 70 75
Gly Ser Glu Cys Arg Ser Phe Ala Gly Pro Gly Ala Pro Phe Pro
80 85 90
Met Asn Arg Ser Ser Gly Thr Pro Gly Arg Pro His Pro Gly Ala
95 100 105
Pro Arg Val Ala Ala Ser Leu Phe Leu Gly Thr Phe Phe Ile Ser
110 115 120
Ser Gly Leu Ile Leu Ser Val Ala Gly Phe Phe Tyr Leu Lys Arg
125 130 135
Ser Ser Lys Leu Pro Arg Ala Cys Tyr Arg Arg Asn Lys Ala Pro
140 145 150
Ala Leu Gln Pro Gly Glu Ala Ala Ala Met Ile Pro Pro Pro Gln
155 160 165
Ser Ser Val Arg Lys Pro Arg Tyr Val Arg Arg Glu Arg Pro Leu
170 175 180
Asp Arg Ala Thr Asp Pro Ala Ala Phe Pro Gly Glu Ala Arg Ile
185 190 195
Ser Asn Val

<210> 35
<211> 73
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2461076CD1

<400> 35
Met Lys Leu Pro Leu Ser Leu Leu Phe Leu Arg Thr Leu Gly Phe
1 5 10 15
Tyr Ile Pro Val Lys Gly Asp Leu Ser Ser Gly Cys Glu Asp Lys
20 25 30
Ala Cys Leu Tyr Val Leu Lys Arg Val Thr Thr Asp Lys Val Phe
35 40 45
Phe Asp Pro Phe Lys Ile Tyr Phe Arg Pro Val Ile Pro Gly Leu
50 55 60
Trp Glu Ala Glu Ala Gly Gly Ser Leu Gly Leu Gly Val
65 70

<210> 36

<211> 376
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1957517CD1

<400> 36
Met Asp Gly Glu Glu Gln Gln Pro Pro His Glu Ala Asn Val Glu
1 5 10 15
Pro Val Val Pro Ser Glu Ala Ser Glu Pro Val Pro Arg Val Leu
20 25 30
Ser Gly Asp Pro Gln Asn Leu Ser Asp Val Asp Ala Phe Asn Leu
35 40 45
Leu Leu Glu Met Lys Leu Lys Arg Arg Arg Gln Arg Pro Asn Leu
50 55 60
Pro Arg Thr Val Thr Gln Leu Val Ala Glu Asp Gly Ser Arg Val
65 70 75
Tyr Val Val Gly Thr Ala His Phe Ser Asp Asp Ser Lys Arg Asp
80 85 90
Val Val Lys Thr Ile Arg Glu Val Gln Pro Asp Val Val Val Val
95 100 105
Glu Leu Cys Gln Tyr Arg Val Ser Met Leu Lys Met Asp Glu Ser
110 115 120
Thr Leu Leu Arg Glu Ala Gln Glu Leu Ser Leu Glu Lys Leu Gln
125 130 135
Gln Ala Val Arg Gln Asn Gly Leu Met Ser Gly Leu Met Gln Met
140 145 150
Leu Leu Leu Lys Val Ser Ala His Ile Thr Glu Gln Leu Gly Met
155 160 165
Ala Pro Gly Gly Glu Phe Arg Glu Ala Phe Lys Glu Ala Ser Lys
170 175 180
Val Pro Phe Cys Lys Phe His Leu Gly Asp Arg Pro Ile Pro Val
185 190 195
Thr Phe Lys Arg Ala Ile Ala Ala Leu Ser Phe Trp Gln Lys Val
200 205 210
Arg Leu Ala Trp Gly Leu Cys Phe Leu Ser Asp Pro Ile Ser Lys
215 220 225
Asp Asp Val Glu Arg Cys Lys Gln Lys Asp Leu Leu Glu Gln Met
230 235 240
Met Ala Glu Met Ile Gly Glu Phe Pro Asp Leu His Arg Thr Ile
245 250 255
Val Ser Glu Arg Asp Val Tyr Leu Thr Tyr Met Leu Arg Gln Ala
260 265 270
Ala Arg Arg Leu Glu Leu Pro Arg Ala Ser Asp Ala Glu Pro Arg
275 280 285
Lys Cys Val Pro Ser Val Val Val Gly Val Val Gly Met Gly His
290 295 300
Val Pro Gly Ile Glu Lys Asn Trp Ser Thr Asp Leu Asn Ile Gln
305 310 315
Glu Ile Met Thr Val Pro Pro Pro Ser Val Ser Gly Arg Val Ser
320 325 330
Arg Leu Ala Val Lys Ala Ala Phe Phe Gly Leu Leu Gly Tyr Ser
335 340 345
Leu Tyr Trp Met Gly Arg Arg Thr Ala Ser Leu Val Leu Ser Leu
350 355 360
Pro Ala Ala Gln Tyr Cys Leu Gln Arg Val Thr Glu Ala Arg His
365 370 375
Lys

<210> 37

<211> 216
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 866038CD1

<400> 37
Met Met Tyr Trp Ile Val Phe Ala Phe Phe Thr Thr Ala Glu Thr
1 5 10 15
Leu Thr Asp Ile Val Leu Ser Trp Phe Pro Phe Tyr Phe Glu Leu
20 25 30
Lys Ile Ala Phe Val Ile Trp Leu Leu Ser Pro Tyr Thr Lys Gly
35 40 45
Ser Ser Val Leu Tyr Arg Lys Phe Val His Pro Thr Leu Ser Asn
50 55 60
Lys Glu Lys Glu Ile Asp Glu Tyr Ile Thr Gln Ala Arg Asp Lys
65 70 75
Ser Tyr Glu Thr Met Met Arg Val Gly Lys Arg Gly Leu Asn Leu
80 85 90
Ala Ala Asn Ala Ala Val Thr Ala Ala Ala Lys Gly Gln Gly Val
95 100 105
Leu Ser Glu Lys Leu Arg Ser Phe Ser Met Gln Asp Leu Thr Leu
110 115 120
Ile Arg Asp Glu Asp Ala Leu Pro Leu Gln Arg Pro Asp Gly Arg
125 130 135
Leu Arg Pro Ser Pro Gly Ser Leu Leu Asp Thr Ile Glu Asp Leu
140 145 150
Gly Asp Asp Pro Ala Leu Ser Leu Arg Ser Ser Thr Asn Pro Ala
155 160 165
Asp Ser Arg Thr Glu Ala Ser Glu Asp Asp Met Gly Asp Lys Ala
170 175 180
Pro Lys Arg Ala Lys Pro Ile Lys Lys Ala Pro Lys Ala Glu Pro
185 190 195
Leu Ala Ser Lys Thr Leu Lys Thr Arg Pro Lys Lys Lys Thr Ser
200 205 210
Gly Gly Gly Asp Ser Ala
215

<210> 38
<211> 233
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3869704CD1

<400> 38
Met Ala Trp Thr Pro Leu Leu Leu Pro Leu Leu Thr Phe Cys Thr
1 5 10 15
Val Ser Glu Ala Ser Tyr Glu Leu Thr Gln Pro Pro Ser Val Ser
20 25 30
Val Ser Pro Gly Gln Thr Ala Arg Ile Thr Cys Ser Gly Asp Ala
35 40 45
Leu Pro Lys Lys Tyr Ala Tyr Trp Tyr Gln Gln Lys Ser Gly Gln
50 55 60
Ala Pro Val Leu Val Ile Tyr Glu Asp Asn Lys Arg Pro Ser Gly
65 70 75
Ile Pro Glu Arg Phe Phe Gly Ser Ser Ser Gly Thr Met Ala Thr
80 85 90
Leu Thr Ile Ser Gly Ala Gln Val Glu Asp Glu Ala Asp Tyr Tyr

```
                 95                  100                 105
Cys Tyr Ser Thr Asp Ser Ser Gly Asn Asp Arg Val Phe Gly Gly
                110                 115                 120
Gly Thr Lys Leu Thr Val Leu Gly Gln Pro Lys Ala Ala Pro Ser
                125                 130                 135
Val Thr Leu Phe Pro Pro Ser Ser Glu Glu Leu Gln Ala Asn Lys
                140                 145                 150
Ala Thr Leu Val Cys Leu Ile Ser Asp Phe Tyr Pro Gly Ala Val
                155                 160                 165
Thr Val Ala Trp Lys Ala Asp Ser Ser Pro Val Lys Ala Gly Val
                170                 175                 180
Glu Thr Thr Thr Pro Ser Lys Gln Ser Asn Asn Lys Tyr Ala Ala
                185                 190                 195
Ser Ser Tyr Leu Ser Leu Thr Pro Glu Gln Trp Lys Ser His Lys
                200                 205                 210
Ser Tyr Ser Cys Gln Val Thr His Glu Gly Ser Thr Val Glu Lys
                215                 220                 225
Thr Val Ala Pro Thr Glu Cys Ser
                230
```

<210> 39
<211> 163
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1415179CD1

<400> 39

```
Met Leu Cys Pro Leu Ser His Ala Arg Val Val Arg Gly Ala Gly
  1                 5                  10                  15
Ser Glu Gly Gly Arg Ile Leu Leu Ser Leu Cys Phe Ser Phe Cys
                 20                  25                  30
Pro Ser Gly Leu Ser Cys Trp Cys Ser Arg His Cys Leu Pro Ala
                 35                  40                  45
Leu Ala Pro Arg Cys Ser Pro Gln Pro Tyr Leu Ser Cys Phe Pro
                 50                  55                  60
Gly Ala Thr His Pro Cys Pro Thr Pro Ser Ala Cys Ser His Gly
                 65                  70                  75
Arg Gly Arg Thr His Ser Leu His Thr His Thr Pro Arg Leu His
                 80                  85                  90
Pro Val Ser Ile Tyr Lys His Val Arg Ala Arg Val His Thr Ser
                 95                 100                 105
Arg Phe Ser Thr Ala Tyr Gln Ala Leu Leu Leu Pro Cys Leu Ser
                110                 115                 120
Ala Trp Arg Gly Pro Pro Leu Leu Thr Pro Ser Val Pro Pro Pro
                125                 130                 135
Glu Leu Ile Arg Met Arg Met Val Val Pro Ala Ser Glu Gly Leu
                140                 145                 150
Leu Gly Leu Leu Gly Ala Lys Pro Leu Cys Pro Lys Gln
                155                 160
```

<210> 40
<211> 235
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1664792CD1

<400> 40

```
Met Arg Leu Lys Leu Phe Ser Ile Leu Ser Thr Val Leu Leu Arg
  1               5                  10                  15
Ala Thr Asp Thr Ile Asn Ser Gln Gly Gln Phe Pro Ser Tyr Leu
                20                  25                  30
Glu Thr Val Thr Lys Asp Ile Leu Ala Pro Asn Leu Gln Trp His
                35                  40                  45
Ala Gly Arg Thr Ala Ala Ala Ile Arg Thr Ala Ala Val Ser Cys
                50                  55                  60
Leu Trp Ala Leu Thr Ser Ser Glu Val Leu Ser Ala Glu Gln Ile
                65                  70                  75
Arg Asp Val Gln Glu Thr Leu Met Pro Gln Val Leu Thr Thr Leu
                80                  85                  90
Glu Glu Asp Ser Lys Met Thr Arg Leu Ile Ser Cys Arg Ile Ile
                95                 100                 105
Asn Thr Phe Leu Lys Thr Ser Gly Gly Met Thr Asp Pro Glu Lys
               110                 115                 120
Leu Ile Lys Ile Tyr Pro Glu Leu Leu Lys Arg Leu Asp Asp Val
               125                 130                 135
Ser Asn Asp Val Arg Met Ala Ala Ala Ser Thr Leu Val Thr Trp
               140                 145                 150
Leu Gln Cys Val Lys Gly Ala Asn Ala Lys Ser Tyr Tyr Gln Ser
               155                 160                 165
Ser Val Gln Tyr Leu Tyr Arg Glu Leu Leu Val His Leu Asp Asp
               170                 175                 180
Pro Glu Arg Ala Ile Gln Asp Ala Ile Leu Glu Val Leu Lys Glu
               185                 190                 195
Gly Ser Gly Leu Phe Pro Asp Leu Leu Val Arg Glu Thr Glu Ala
               200                 205                 210
Val Ile His Lys His Arg Ser Ala Thr Tyr Cys Glu Gln Leu Leu
               215                 220                 225
Gln His Val Gln Ala Val Pro Ala Thr Gln
               230                 235
```

<210> 41
<211> 94
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2079396CD1

<400> 41

```
Met Ser Pro Leu Ser Pro Thr Gly Leu Asn Leu Trp Gly Gly Glu
  1               5                  10                  15
Gly Ser Ser Leu His Ser Ala Leu Asp His Gln Gly Arg Gly Ile
                20                  25                  30
Thr Leu Ala Ile Gly Ile Ile Ser Ser Ser Phe Ser Ser Pro Ser
                35                  40                  45
Pro Arg Ile Arg Pro Ser Ser Gln His Cys Val Gly Leu Ile Leu
                50                  55                  60
Arg Ile Leu Tyr His His Pro Gly Leu Gly Gly Cys Arg Ser Trp
                65                  70                  75
Val Leu Leu Leu Arg Asp Arg Val Ser Leu Cys His Pro Gly Trp
                80                  85                  90
Ser Ala Val Ala
```

<210> 42
<211> 85
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 5390115CD1

<400> 42
Met Ala Ser Asp Leu Asp Phe Ser Pro Pro Glu Val Pro Glu Pro
1 5 10 15
Thr Phe Leu Glu Asn Leu Leu Arg Tyr Gly Leu Phe Leu Gly Ala
20 25 30
Ile Phe Gln Leu Ile Cys Val Leu Ala Ile Ile Val Pro Ile Pro
35 40 45
Lys Ser His Glu Ala Glu Ala Glu Pro Ser Glu Pro Arg Ser Ala
50 55 60
Glu Val Thr Arg Lys Pro Lys Ala Ala Val Pro Ser Val Asn Lys
65 70 75
Arg Pro Lys Lys Glu Thr Lys Lys Lys Arg
80 85

<210> 43
<211> 901
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1403326CD1

<400> 43
Met Glu Ala Asn Gln Cys Pro Leu Val Val Glu Pro Ser Tyr Pro
1 5 10 15
Asp Leu Val Ile Asn Val Gly Glu Val Thr Leu Gly Glu Glu Asn
20 25 30
Arg Lys Lys Leu Gln Lys Ile Gln Arg Asp Gln Glu Lys Glu Arg
35 40 45
Val Met Arg Ala Ala Cys Ala Leu Leu Asn Ser Gly Gly Gly Val
50 55 60
Ile Arg Met Ala Lys Lys Val Glu His Pro Val Glu Met Gly Leu
65 70 75
Asp Leu Glu Gln Ser Leu Arg Glu Leu Ile Gln Ser Ser Asp Leu
80 85 90
Gln Ala Phe Phe Glu Thr Lys Gln Gln Gly Arg Cys Phe Tyr Ile
95 100 105
Phe Val Lys Ser Trp Ser Ser Gly Pro Phe Pro Glu Asp Arg Ser
110 115 120
Phe Lys Pro Arg Leu Cys Ser Leu Ser Ser Ser Leu Tyr Arg Arg
125 130 135
Ser Glu Thr Ser Val Arg Ser Met Asp Ser Arg Glu Ala Phe Cys
140 145 150
Phe Leu Lys Thr Lys Arg Lys Pro Lys Ile Leu Glu Glu Gly Pro
155 160 165
Phe His Lys Ile His Lys Gly Val Tyr Gln Glu Leu Pro Asn Ser
170 175 180
Asp Pro Ala Asp Pro Asn Ser Asp Pro Ala Asp Leu Ile Phe Gln
185 190 195
Lys Asp Tyr Leu Glu Tyr Gly Glu Ile Leu Pro Phe Pro Glu Ser
200 205 210
Gln Leu Val Glu Phe Lys Gln Phe Ser Thr Lys His Phe Gln Glu
215 220 225
Tyr Val Lys Arg Thr Ile Pro Glu Tyr Val Pro Ala Phe Ala Asn
230 235 240
Thr Gly Gly Gly Tyr Leu Phe Ile Gly Val Asp Asp Lys Ser Arg
245 250 255
Glu Val Leu Gly Cys Ala Lys Glu Asn Val Asp Pro Asp Ser Leu

```
                260                 265                 270
Arg Arg Lys Ile Glu Gln Ala Ile Tyr Lys Leu Pro Cys Val His
                275                 280                 285
Phe Cys Gln Pro Gln Arg Pro Ile Thr Phe Thr Leu Lys Ile Val
                290                 295                 300
Asp Val Leu Lys Arg Gly Glu Leu Tyr Gly Tyr Ala Cys Met Ile
                305                 310                 315
Arg Val Asn Pro Phe Cys Cys Ala Val Phe Ser Glu Ala Pro Asn
                320                 325                 330
Ser Trp Ile Val Glu Asp Lys Tyr Val Cys Ser Leu Thr Thr Glu
                335                 340                 345
Lys Trp Val Gly Met Met Thr Asp Thr Asp Pro Asp Leu Leu Gln
                350                 355                 360
Leu Ser Glu Asp Phe Glu Cys Gln Leu Ser Leu Ser Ser Gly Pro
                365                 370                 375
Pro Leu Ser Arg Pro Val Tyr Ser Lys Lys Gly Leu Glu His Lys
                380                 385                 390
Ala Asp Leu Gln Gln His Leu Phe Pro Val Pro Pro Gly His Leu
                395                 400                 405
Glu Cys Thr Pro Glu Ser Leu Trp Lys Glu Leu Ser Leu Gln His
                410                 415                 420
Glu Gly Leu Lys Glu Leu Ile His Lys Gln Met Arg Pro Phe Ser
                425                 430                 435
Gln Gly Ile Val Ile Leu Ser Arg Ser Trp Ala Val Asp Leu Asn
                440                 445                 450
Leu Gln Glu Lys Pro Gly Val Ile Cys Asp Ala Leu Leu Ile Ala
                455                 460                 465
Gln Asn Ser Thr Pro Ile Leu Tyr Thr Ile Leu Arg Glu Gln Asp
                470                 475                 480
Ala Glu Gly Gln Asp Tyr Cys Thr Arg Thr Ala Phe Thr Leu Lys
                485                 490                 495
Gln Lys Leu Val Asn Met Gly Gly Tyr Thr Gly Lys Val Cys Val
                500                 505                 510
Arg Ala Lys Val Leu Cys Leu Ser Pro Glu Ser Ser Ala Glu Ala
                515                 520                 525
Leu Glu Ala Ala Val Ser Pro Met Asp Tyr Pro Ala Ser Tyr Ser
                530                 535                 540
Leu Ala Gly Thr Gln His Met Glu Ala Leu Leu Gln Ser Leu Val
                545                 550                 555
Ile Val Leu Leu Gly Phe Arg Ser Leu Leu Ser Asp Gln Leu Gly
                560                 565                 570
Cys Glu Val Leu Asn Leu Leu Thr Ala Gln Gln Tyr Glu Ile Phe
                575                 580                 585
Ser Arg Ser Leu Arg Lys Asn Arg Glu Leu Phe Val His Gly Leu
                590                 595                 600
Pro Gly Ser Gly Lys Thr Ile Met Ala Met Lys Ile Met Glu Lys
                605                 610                 615
Ile Arg Asn Val Phe His Cys Glu Ala His Arg Ile Leu Tyr Val
                620                 625                 630
Cys Glu Asn Gln Pro Leu Arg Asn Phe Ile Ser Asp Arg Asn Ile
                635                 640                 645
Cys Arg Ala Glu Thr Arg Lys Thr Phe Leu Arg Glu Asn Phe Glu
                650                 655                 660
His Ile Gln His Ile Val Ile Asp Glu Ala Gln Asn Phe Arg Thr
                665                 670                 675
Glu Asp Gly Asp Trp Tyr Gly Lys Ala Lys Ser Ile Thr Arg Arg
                680                 685                 690
Ala Lys Gly Gly Pro Gly Ile Leu Trp Ile Phe Leu Asp Tyr Phe
                695                 700                 705
Gln Thr Ser His Leu Asp Cys Ser Gly Leu Pro Pro Leu Ser Asp
                710                 715                 720
Gln Tyr Pro Arg Glu Glu Leu Thr Arg Ile Val Arg Asn Ala Asp
                725                 730                 735
```

```
Pro Ile Ala Lys Tyr Leu Gln Lys Glu Met Gln Val Ile Arg Ser
                740                 745                 750
Asn Pro Ser Phe Asn Ile Pro Thr Gly Cys Leu Glu Val Phe Pro
                755                 760                 765
Glu Ala Glu Trp Ser Gln Gly Val Gln Gly Thr Leu Arg Ile Lys
                770                 775                 780
Lys Tyr Leu Thr Val Glu Gln Ile Met Thr Cys Val Ala Asp Thr
                785                 790                 795
Cys Arg Arg Phe Phe Asp Arg Gly Tyr Ser Pro Lys Asp Val Ala
                800                 805                 810
Val Leu Val Ser Thr Ala Lys Glu Val Glu His Tyr Lys Tyr Glu
                815                 820                 825
Leu Leu Lys Ala Met Arg Lys Lys Arg Val Val Gln Leu Ser Asp
                830                 835                 840
Ala Cys Asp Met Leu Gly Asp His Ile Val Leu Asp Ser Val Arg
                845                 850                 855
Arg Phe Ser Gly Leu Glu Arg Ser Ile Val Phe Gly Ile His Pro
                860                 865                 870
Arg Thr Ala Asp Pro Ala Ile Leu Pro Asn Val Leu Ile Cys Leu
                875                 880                 885
Ala Ser Arg Ala Lys Gln His Leu Tyr Ile Phe Pro Trp Gly Gly
                890                 895                 900
His
```

<210> 44
<211> 1040
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7690129CD1

<400> 44

```
Met Ala Ser Thr Gly Gly Thr Lys Val Val Ala Met Gly Val Ala
  1               5                  10                  15
Pro Trp Gly Val Val Arg Asn Arg Asp Thr Leu Ile Asn Pro Lys
                20                  25                  30
Gly Ser Phe Pro Ala Arg Tyr Arg Trp Arg Gly Asp Pro Glu Asp
                35                  40                  45
Gly Val Gln Phe Pro Leu Asp Tyr Asn Tyr Ser Ala Phe Phe Leu
                50                  55                  60
Val Asp Asp Gly Thr His Gly Cys Leu Gly Gly Glu Asn Arg Phe
                65                  70                  75
Arg Leu Arg Leu Glu Ser Tyr Ile Ser Gln Gln Lys Thr Gly Val
                80                  85                  90
Gly Gly Thr Gly Ile Asp Ile Pro Val Leu Leu Leu Leu Ile Asp
                95                 100                 105
Gly Asp Glu Lys Met Leu Thr Arg Ile Glu Asn Ala Thr Gln Ala
               110                 115                 120
Gln Leu Pro Cys Leu Leu Val Ala Gly Ser Gly Gly Ala Ala Asp
               125                 130                 135
Cys Leu Ala Glu Thr Leu Glu Asp Thr Leu Ala Pro Gly Ser Gly
               140                 145                 150
Gly Ala Arg Gln Gly Glu Ala Arg Asp Arg Ile Arg Arg Phe Phe
               155                 160                 165
Pro Lys Gly Asp Leu Glu Val Leu Gln Ala Gln Val Glu Arg Ile
               170                 175                 180
Met Thr Arg Lys Glu Leu Leu Thr Val Tyr Ser Ser Glu Asp Gly
               185                 190                 195
Ser Glu Glu Phe Glu Thr Ile Val Leu Lys Ala Leu Val Lys Ala
               200                 205                 210
```

```
Cys Gly Ser Ser Glu Ala Ser Ala Tyr Leu Asp Glu Leu Arg Leu
                215                 220                 225
Ala Val Ala Trp Asn Arg Val Asp Ile Ala Gln Ser Glu Leu Phe
                230                 235                 240
Arg Gly Asp Ile Gln Trp Arg Ser Phe His Leu Glu Ala Ser Leu
                245                 250                 255
Met Asp Ala Leu Leu Asn Asp Arg Pro Glu Phe Val Arg Leu Leu
                260                 265                 270
Ile Ser His Gly Leu Ser Leu Gly His Phe Leu Thr Pro Met Arg
                275                 280                 285
Leu Ala Gln Leu Tyr Ser Ala Ala Pro Ser Asn Ser Leu Ile Arg
                290                 295                 300
Asn Leu Leu Asp Gln Ala Ser His Ser Ala Gly Thr Lys Ala Pro
                305                 310                 315
Ala Leu Lys Gly Gly Ala Ala Glu Leu Arg Pro Pro Asp Val Gly
                320                 325                 330
His Val Leu Arg Met Leu Leu Gly Lys Met Cys Ala Pro Arg Tyr
                335                 340                 345
Pro Ser Gly Gly Ala Trp Asp Pro His Pro Gly Gln Gly Phe Gly
                350                 355                 360
Glu Ser Met Tyr Leu Leu Ser Asp Lys Ala Thr Ser Pro Leu Ser
                365                 370                 375
Leu Asp Ala Gly Leu Gly Gln Ala Pro Trp Ser Asp Leu Leu Leu
                380                 385                 390
Trp Ala Leu Leu Leu Asn Arg Ala Gln Met Ala Met Tyr Phe Trp
                395                 400                 405
Glu Met Gly Ser Asn Ala Val Ser Ser Ala Leu Gly Ala Cys Leu
                410                 415                 420
Leu Leu Arg Val Met Ala Arg Leu Glu Pro Asp Ala Glu Glu Ala
                425                 430                 435
Ala Arg Arg Lys Asp Leu Ala Phe Lys Phe Glu Gly Met Gly Val
                440                 445                 450
Asp Leu Phe Gly Glu Cys Tyr Arg Ser Ser Glu Val Arg Ala Ala
                455                 460                 465
Arg Leu Leu Leu Arg Arg Cys Pro Leu Trp Gly Asp Ala Thr Cys
                470                 475                 480
Leu Gln Leu Ala Met Gln Ala Asp Ala Arg Ala Phe Phe Ala Gln
                485                 490                 495
Asp Gly Val Gln Ser Leu Leu Thr Gln Lys Trp Trp Gly Asp Met
                500                 505                 510
Ala Ser Thr Thr Pro Ile Trp Ala Leu Val Leu Ala Phe Phe Cys
                515                 520                 525
Pro Pro Leu Ile Tyr Thr Arg Leu Ile Thr Phe Arg Lys Ser Glu
                530                 535                 540
Glu Glu Pro Thr Arg Glu Glu Leu Glu Phe Asp Met Asp Ser Val
                545                 550                 555
Ile Asn Gly Glu Gly Pro Val Gly Thr Ala Asp Pro Ala Glu Lys
                560                 565                 570
Thr Pro Leu Gly Val Pro Arg Gln Ser Gly Arg Pro Gly Cys Cys
                575                 580                 585
Gly Gly Arg Cys Gly Gly Arg Arg Cys Leu Arg Arg Trp Phe His
                590                 595                 600
Phe Trp Gly Ala Pro Val Thr Ile Phe Met Gly Asn Val Val Ser
                605                 610                 615
Tyr Leu Leu Phe Leu Leu Leu Phe Ser Arg Val Leu Leu Val Asp
                620                 625                 630
Phe Gln Pro Ala Pro Pro Gly Ser Leu Glu Leu Leu Leu Tyr Phe
                635                 640                 645
Trp Ala Phe Thr Leu Leu Cys Glu Glu Leu Arg Gln Gly Leu Ser
                650                 655                 660
Gly Gly Gly Gly Ser Leu Ala Ser Gly Gly Pro Gly Pro Gly His
                665                 670                 675
Ala Ser Leu Ser Gln Arg Leu Arg Leu Tyr Leu Ala Asp Ser Trp
```

```
                680                 685                 690
Asn Gln Cys Asp Leu Val Ala Leu Thr Cys Phe Leu Leu Gly Val
                695                 700                 705
Gly Cys Arg Leu Thr Pro Gly Leu Tyr His Leu Gly Arg Thr Val
                710                 715                 720
Leu Cys Ile Asp Phe Met Val Phe Thr Val Arg Leu Leu His Ile
                725                 730                 735
Phe Thr Val Asn Lys Gln Leu Gly Pro Lys Ile Val Ile Val Ser
                740                 745                 750
Lys Met Met Lys Asp Val Phe Phe Phe Leu Phe Phe Leu Gly Val
                755                 760                 765
Trp Leu Val Ala Tyr Gly Val Ala Thr Glu Gly Leu Leu Arg Pro
                770                 775                 780
Arg Asp Ser Asp Phe Pro Ser Ile Leu Arg Arg Val Phe Tyr Arg
                785                 790                 795
Pro Tyr Leu Gln Ile Phe Gly Gln Ile Pro Gln Glu Asp Met Asp
                800                 805                 810
Val Ala Leu Met Glu His Ser Asn Cys Ser Ser Glu Pro Gly Phe
                815                 820                 825
Trp Ala His Pro Pro Gly Ala Gln Ala Gly Thr Cys Val Ser Gln
                830                 835                 840
Tyr Ala Asn Trp Leu Val Val Leu Leu Leu Val Ile Phe Leu Leu
                845                 850                 855
Val Ala Asn Ile Leu Leu Val Asn Leu Leu Ile Ala Met Phe Ser
                860                 865                 870
Tyr Thr Phe Gly Lys Val Gln Gly Asn Ser Asp Leu Tyr Trp Lys
                875                 880                 885
Ala Gln Arg Tyr Arg Leu Ile Arg Glu Phe His Ser Arg Pro Ala
                890                 895                 900
Leu Ala Pro Pro Phe Ile Val Ile Ser His Leu Arg Leu Leu Leu
                905                 910                 915
Arg Gln Leu Cys Arg Arg Pro Arg Ser Pro Gln Pro Ser Ser Pro
                920                 925                 930
Ala Leu Glu His Phe Arg Val Tyr Leu Ser Lys Glu Ala Glu Arg
                935                 940                 945
Lys Leu Leu Thr Trp Glu Ser Val His Lys Glu Asn Phe Leu Leu
                950                 955                 960
Ala Arg Ala Arg Asp Lys Arg Glu Ser Asp Ser Glu Arg Leu Lys
                965                 970                 975
Arg Thr Ser Gln Lys Val Asp Leu Ala Leu Lys Gln Leu Gly His
                980                 985                 990
Ile Arg Glu Tyr Glu Gln Arg Leu Lys Val Leu Glu Arg Glu Val
                995                1000                1005
Gln Gln Cys Ser Arg Val Leu Gly Trp Val Ala Glu Ala Leu Ser
               1010                1015                1020
Arg Ser Ala Leu Leu Pro Pro Gly Gly Pro Pro Pro Pro Asp Leu
               1025                1030                1035
Pro Gly Ser Lys Asp
               1040
```

<210> 45
<211> 2508
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2101688CB1

<400> 45
cgtgctcctc ccggggtgct tggcacagcc tcggattcct ccctctcgct gctcgagtca 60
gtttccctat cggcggcagc gggcaaggcg gcggcggcgg cggcggcagc cgcggtggcg 120
gcgtggggaa catctcggca gccaccgcgc ttctcccgct ggagcgggcg tccagcttgg 180

ctgccctcgg tccttccctg ccacgtttcg ggtcgccctg cacccccac ccaggctcgc 240
ttctcttcga agcgggaagg gcgccttgca ggatcctgcc gcccctccaa ccggatcctg 300
ggtctagagc tccccagagc gaggcgctcg ccaggactcc tgccccgcca accctgaccg 360
ccgggggtg cccccgggac gtagcgccgc ggagaggaag cggcaaaggg gaccatgcgg 420
cgcctgactc gtcggctggt tctgccagtc ttcggggtgc tctggatcac ggtgctgctg 480
ttcttctggg taaccaagag gaagttggag gtgccgacgg gacctgaagt gcagacccct 540
aagccttcgg acgctgactg ggacgacctg tgggaccagt ttgatgagcg gcggtatctg 600
aatgccaaaa agtggcgcgt tggtgacgac ccctataagc tgtatgcttt caaccagcgg 660
gagagtgagc ggatctccag caatcgggcc atcccggaca ctcgccatct gagatgcaca 720
ctgctggtgt attgcacgga ccttccaccc actagcatca tcatcacctt ccacaacgag 780
gcccgctcca cgctgctcag gaccatccgc agtgtattaa accgcacccc tacgcatctg 840
atccgggaaa tcatattagt ggatgacttc agcaatgacc ctgatgactg taaacagctc 900
atcaagttgc ccaaggtgaa atgcttgcgc aataatgaac ggcaaggtct ggtccggtcc 960
cggattcggg gcgctgacat cgcccagggc accactctga ctttcctcga cagccactgt 1020
gaggtgaaca gggactggct ccagcctctg ttgcacaggg tcaaagagga ctacacgcgg 1080
gtggtgtgcc ctgtgatcga tatcattaac ctggacacct tcacctacat cgagtctgcc 1140
tcggagctca gaggggggtt tgactggagc ctccacttcc agtgggagca gctctcccca 1200
gagcagaagg ctcggcgcct ggaccccacg gagcccatca ggactcctat catagctgga 1260
gggctcttcg tgatcgacaa agcttggttt gattacctgg ggaaatatga tatggacatg 1320
gacatctggg gtggggagaa ctttgaaatc tccttccgag tgtggatgtg cgggggcagc 1380
ctagagatcg tcccctgcag ccgagtgggg cacgtcttcc ggaagaagca cccctacgtt 1440
ttccctgatg gaaatgccaa cacgtatata aagaacacca agcggacagc tgaagtgtgg 1500
atggatgaat acaagcaata ctattacgct gcccggccat tcgccctgga gaggcccttc 1560
gggaatgttg agagcagatt ggacctgagg aagaatctgc gctgccagag cttcaagtgg 1620
tacctggaga atatctaccc tgaactcagc atccccaagg agtcctccat ccagaagggc 1680
aatatccgac agagacagaa gtgcctggaa tctcaaaggc agaacaacca agaaacccca 1740
aacctaaagt tgagcccctg tgccaaggtc aaaggcgaag atgcaaagtc ccaggtatgg 1800
gccttcacat acacccagaa gatcctccag gaggagctgt gcctgtcagt catcaccttg 1860
ttccctggcg ccccagtggt tcttgtcctt tgcaagaatg gagatgaccg acagcaatgg 1920
accaaaactg gttcccacat cgagcacata gcatcccacc tctgcctcga tacagatatg 1980
ttcggtgatg gcaccgagaa cggcaaggaa atcgtcgtca acccatgtga gtcctcactc 2040
atgagccagc actgggacat ggtgagctct tgaggacccc tgccagaagc agcaagggcc 2100
atggggtggt gcttccctgg accagaacag actggaaact gggcagcaag cagcctgcaa 2160
ccacctcaga catcctggac tgggaggtgg aggcagagcc ccccaggaca ggagcaactg 2220
tctcagggag gacagaggaa aacatcacaa gccaatgggg ctcaaagaca aatcccacat 2280
gttctcaagg ccgttaagtt ccagtcctgg ccagtcattc cctgattggt atctggagac 2340
agaaacctaa tgggaagtgt ttattgttcc ttttcctaca aaggaagcag tctctggagg 2400
ccagaaagaa aagccttctt tttcactagg ccaggactac attgagagat gaagaatgga 2460
ggttgtttcc aaaagaaata aagagaaact tagaaaaaaa aaaaaaaa 2508

<210> 46
<211> 4034
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 5452330CB1

<400> 46
aaccagcacc atgcgcccgg tagccctgct gctcctgccc tcgctgctgg cgctcctggc 60
tcacggactc tctttagagg ccccaaccgt ggggaaagga caagccccag gcatcgagga 120
gacagatggc gagctgacag cagcccccac acctgagcag ccagaacgag gcgtccactt 180
tgtcacaaca gcccccacct tgaagctgct caaccaccac ccgctgcttg aggaattcct 240
acaagagggg ctggaaaagg gagatgagga gctgaggcca gcactgccct tccagcctga 300
cccacctgca cccttcaccc caagtccctt tccccgcctg gccaaccagg acagccgccc 360
tgtctttacc agccccactc cagccatggc tgcggtaccc actcagcccc agtccaagga 420
gggaccctgg agtccggagt cagagtcccc tatgcttcga atcacagctc ccctacctcc 480
agggcccagc atggcagtgc ccaccctagg cccaggggag atagccagca ctacaccccc 540
cagcagagcc tggacaccaa cccaagaggg tcctggagac atgggaaggc cgtgggttgc 600
agaggttgtg tcccagggcg cagggatcgg gatccagggg accatcacct cctccacagc 660
ttcaggagat gatgaggaga ccaccactac caccaccatc atcaccacca ccatcaccac 720
agtccagaca ccaggccctt gtagctggaa tttctcaggc ccagagggct ctctggactc 780

ccctacagac ctcagctccc ccactgatgt tggcctggac tgcttcttct acatctctgt 840
ctaccctggc tatggcgtgg aaatcaaggt ccagaatatc agcctccggg aaggggagac 900
agtgactgtg gaaggcctgg gggggcctga cccactgccc ctggccaacc agtctttcct 960
gctgcggggc caagtcatcc gcagccccac ccaccaagcg gccctgaggt tccagagcct 1020
cccgccaccg gctggccctg gcaccttcca tttccattac caagcctatc tcctgagctg 1080
ccactttccc cgtcgtccag cttatggaga tgtgactgtc accagcctcc acccaggggg 1140
tagtgcccgc ttccattgtg ccactggcta ccagctgaag ggcgccaggc atctcacctg 1200
tctcaatgcc acccagccct tctgggattc aaaggagccc gtctgcatcg ctgcttgcgg 1260
cggagtgatc cgcaatgcca ccaccggccg catcgtctct ccaggcttcc cgggcaacta 1320
cagcaacaac ctcacctgtc actggctgct tgaggctcct gagggccagc ggctacacct 1380
gcactttgag aaggtttccc tggcagagga tgatgacagg ctcatcattc gcaatgggga 1440
caacgtggag gccccaccag tgtatgattc ctatgaggtg gaatacctgc ccattgaggg 1500
cctgctcagc tctggcaaac acttctttgt tgagctcagt actgacagca gcggggcagc 1560
tgcaggcatg gccctgcgct atgaggcctt ccagcagggc cattgctatg agccctttgt 1620
caaatacggt aacttcagca gcagcacacc cacctaccct gtgggtacca ctgtggagtt 1680
cagctgcgac cctggctaca ccctggagca gggctccatc atcatcgagt gtgttgaccc 1740
ccacgacccc cagtggaatg agacagagcc agcctgccga gccgtgtgca gcggggagat 1800
cacagactcg gctggcgtgg tactctctcc caactggcca gagccctacg gtcgtgggca 1860
ggattgtatc tggggtgtgc atgtggaaga ggacaagcgc atcatgctgg acatccgagt 1920
gctgcgcata ggccctggtg atgtgcttac cttctatgat ggggatgacc tgacggcccg 1980
ggttctgggc cagtactcag ggccccgtag ccacttcaag ctctttacct ccatggctga 2040
tgtcaccatt cagttccagt cggaccccgg gacctcagtg ctgggctacc agcagggctt 2100
cgtcatccac ttctttgagg tgccccgcaa tgacacatgt ccggagctgc ctgagatccc 2160
caatggctgg aagagcccat cgcagcctga gctagtgcac ggcaccgtgg tcacttacca 2220
gtgctaccct ggctaccagg tagtgggatc cagtgtcctc atgtgccagt gggacctaac 2280
ttggagtgag gacctgccct catgccagag ggtgacttcc tgccacgatc ctggagatgt 2340
ggagcacagc cgacgcctca tatccagccc caagtttccc gtgggggcca ccgtgcaata 2400
tatctgtgac cagggttttg tgctgatggg cagctccatc ctcacctgcc atgatcgcca 2460
ggctggcagc cccaagtgga gtgaccgggc ccctaaatgt ctcctggaac agctcaagcc 2520
atgccatggt ctcagtgccc ctgagaatgg tgcccgaagt cctgagaagc agctacaccc 2580
agcaggggcc accatccact tctcgtgtgc ccctggctat gtgctgaagg gccaggccag 2640
catcaagtgt gtgcctgggc acccctcgca ttggagtgac cccccaccca tctgtagggc 2700
tgcctctctg gatgggttct acaacagtcg cagcctggat gttgccaagg cacctgctgc 2760
ctccagcacc ctggatgctg cccacattgc agctgccatc ttcttgccac tggtggcgat 2820
ggtgttgttg gtaggaggtg tatacttcta cttctccagg ctccagggaa aaagctccct 2880
gcagctgccc cgcccccgcc cccgccccta caaccgcatt accatagagt cagcgtttga 2940
caatccaact tacgagactg gatctctttc ctttgcagga gacgagagaa tatgaagtct 3000
ccatctaggt gggggcagtc tagggaagtc aactcagact tgcaccacag tccagcagca 3060
aggctccttg cttcctgctg tccctccacc tcctgtatat accacctagg aggagatgcc 3120
accaagccct caagaagttg tgcccttccc cgcctgcgat gcccaccatg gcctattttc 3180
ttggtgtcat tgcccacttg gggcccttca ttgggcccat gtcagggggc atctacctgt 3240
gggaagaaca tagctggagc acaagcatca acagccagca tcctgagcct cctcatgccc 3300
tggaccagcc tggaacacac tagcagagca ggagtacctt tctccacatg accaccatcc 3360
cgccctggca tggcaacctg cagcaggatt aacttgacca tggtgggaac tgcaccaggg 3420
tactcctcac agcgccatca ccaatggcca aaactcctct caacggtgac ctctgggtag 3480
tcctggcatg ccaacatcag cctcttggga ggtctctagt tctctaaagt tctggacagt 3540
tctgcctcct gccctgtccc agtggaggca gtaattctag gagatcctaa ggggttcagg 3600
gggaccctac ccccacctca ggttgggctt ccctgggcac tcatgctcca caccaaagca 3660
ggacacgcca ttttccactg accaccctat accctgagga aagggagact ttcctccgat 3720
gtttatttag ctgttgcaaa catcttcacc ctaatagtcc ctcctccaat tccagccact 3780
tgtcaggctc tcctcttgac cactgtgtta tgggataagg ggaggggtg ggcatattct 3840
ggagaggagc agaggtccaa ggacccagga atttggcatg gaacaggtgg taggagagcc 3900
ccagggagac gcccaggagc tggctgaaag ccactttgta catgtaatgt attatatggg 3960
gtctgggctc cagccagaga acaatctttt atttctgttg tttccttatt aaaatggtgt 4020
ttttggaaaa aaaa 4034

<210> 47
<211> 845
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature

<223> Incyte ID No: 4362432CB1

<400> 47
ggtagaaaag ccagcctatg ttacaggaca aaggccgtcg ctttgtaaaa gcttgaagtg 60
cagtttgctg ctgagtacag aagacctttg caaacagaga ggggagattt tctctgtaag 120
gttgcaaaca agagcaggtc ctggaagata agattccccg ccatgttatc ctccgtggtg 180
ttttggggac taattgccct cattggcact tccaggggct catacccctt cagtcactca 240
atgaagcctc acctacatcc acgcctgtac cacggctgct atggggacat catgaccatg 300
aagacctctg gggccacttg tgatgcaaac agtgtgatga actgcgggat ccgtggttct 360
gaaatgtttg ctgagatgga tttgagggcc ataaaacctt accagactct gatcaaagaa 420
gtcgggcaga gacattgcgt ggaccctgct gtcatcgcag ccatcatctc cagggaaagc 480
catggcggat ctgtcctgca agacggctgg gaccacaggg gacttaaatt tggcttgatg 540
cagcttgata aacaaacgta ccaccctgtc ggtgcctggg atagcaaaga gcacctttca 600
caggctactg ggattctaac agagagaatt aaggcaatcc agaaaaaatt ccccacgtgg 660
agtgttgctc agcacctcaa aggtggtctc tcagctttta agtcaggaat tgaagcgatt 720
gccaccccat cggacataga caatgacttc gtcaatgata tcattgctcg agctaagttc 780
tataaaagac aaagcttcta ggcaaagctc tgtgggtggg ccaggttggc agagtgctca 840
gatgt 845

<210> 48
<211> 2300
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 5308104CB1

<400> 48
tttacggcgc agtgtgctgg acaagcggtt caggccgggc ggcgttgttg gcgggggccc 60
cggtggaggc ccggcctggg cggcgcccgc catgaatggg ctgtcgctga gtgagctctg 120
ctgcctcttc tgctgcccgc cctgccccgg ccgcatcgct gccaagctcg ccttcctgcc 180
gccggaggcc acctactccc tggtgcctga gcccgagccg gggcctggtg gggccggggc 240
cgcccccttg gggaccctga gagcctcctc gggcgcaccc gggcgctgga agctgcacct 300
gacggagcgt gccgacttcc agtacagcca gcgcgagctg gacaccatcg aggtcttccc 360
caccaagagc gcccgcggca accgcgtctc ctgcatgtat gttcgctgcg tgcctggtgc 420
caggtacacg gtcctcttct cgcacggcaa tgccgtggac ctgggccaga tgagcagctt 480
ctacattggc ctgggctccc gcctccactg caacatcttc tcctacgact actccggcta 540
cggtgccagc tcgggcaggc cttccgagag gaacctctat gccgacatcg acgccgcctg 600
gcaggccctg cgcaccaggt acggcatcag cccggacagc atcatcctgt acgggcagag 660
catcggcacg gtgcccaccg tggacctggc ctcgcgctac gagtgtgccg cggtaattct 720
ccattccccc ctgatgtctg gtttgcgtgt ggcttttccg gataccagga aaacatactg 780
ctttgatgct ttccccagca ttgacaagat atctaaagtc acctctcctg tgttggtcat 840
tcatggtaca gaggatgagg tcatcgattt ctcccatggc ctagcgatgt acgagcgctg 900
tccccgagcc gtggagcccc tttgggttga aggggctggg cataatgaca tagagcttta 960
tgcacaatac ctagaaagac taaaacagtt catatctcac gaacttccta attcctgaag 1020
acaacaactt gatcttacct catttactgt gaacagaaga gtcctctgtt ttgcacatgc 1080
tttaactggg tagctgtaaa ggcttgataa ccatgaagaa gtgcccaacc tttagggtgt 1140
tctaatcaaa gagctgatga aatctcagtc ttttgtatct agaggtggtt ctgctaattc 1200
acacaacacg ttaaactgaa cagtcgtgat tcccagcttc attaccttgc aggaatggga 1260
atgagagctg aatgtaggga caattttcta gtgctgtata aagtagcctc gcatctgttt 1320
ctcaacctta tccatcattt ctgacattca tgcaggactt gccctgttgc caccaatgtt 1380
ctcggtattt cacatgcagc tctctttctg ccactggata catgggttca atccatttgt 1440
gaagctgtga tagtgtaact ggaaagctag tgtggtgaaa attcctttat tatttttttgt 1500
taacatgctg atctttcccg gacaaatgaa ctgaagggta atttactgga actctcgtgt 1560
acagcttcat caactgtaac catataaata taactggaat attcttaaac aaaaagaaac 1620
taggggtttt tttaagtgta aatttattac tagccaacag agttttacta ttttgattgt 1680
ctggttggtt taacaaagag cctagctgac tttccttctg taaagtcctc cttgtaggct 1740
tttttaaagt actgtacata tttgcaatca cattgtgcat agattcttaa tggtagatat 1800
gatttctttt gtcaggctac aacaatgaac tgcagattcc ttgtttgtaa tgtaaatgat 1860
tgaatacatt ttgttaatat gtttttattc ctatgttttg ctattaaaaa ttttataaca 1920
tttccaagac aaaaattcca agtttatgct ttgaagaatt tatgtaatta aaatttcact 1980
aaactaatct ttttagttta ggaattattt gggttttgac actggaagtt gcgccaaata 2040

agcatcagaa ataggagatg cttaacattg ctatactact tgtgttggtt aggggtttgg 2100
atttgggggt tctttggttt taatttttt ttccacattt aaaagcctta aatgtactgt 2160
aagcctcaga tcgttgtaca actggactgc ggttgattgc cagtttgtgt actgttgctt 2220
ggatgcggca cagtggttgg taatggaata aaggatgcat ggatcagaaa aaaaaaaaaa 2280
aaaaaaaaaa aaaaaaaaaa 2300

<210> 49
<211> 1587
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3092736CB1

<400> 49
ggaggacata attcacctgt cctagcttct tatcatctta catttccctg tagccactgg 60
gacatatgtg gtgttccttc ctagctcctg tctcctcctc atgcctttgc tgggtatggg 120
catgttgggg ggaaggtcat tgctgtcaga ggggcactga ctttctaatg gtgttaccca 180
aggtgaatgt tggagacaca gtcgcgatgc tgcccaagtc ccggcgagcc ctaactatcc 240
aggagatcgc tgcgctggcc aggtcctccc tgcatggtat ttcccaggtg gtgaaggacc 300
acgtgaccaa gcctaccgcc atggcccagg gccgagtggc tcacctcatt gagtggaagg 360
gctggagcaa gccgagtgac tcacctgctg ccctggaatc agccttttcc tcctattcag 420
acctcagcga gggcgaacaa gaggctcgct ttgcagcagg agtggctgag cagtttgcca 480
tcgcggaagc caagctccga gcatggtctt cggtggatgg cgaggactcc actgatgact 540
cctatgatga ggactttgct gggggaatgg acacagacat ggctgggcag ctgcccctgg 600
ggccgcacct ccaggacctg ttcaccggcc accggttctc ccggcctgtg cgccagggct 660
ccgtggagcc tgagagcgac tgctcacaga ccgtgtcccc agacaccctg tgctctagtc 720
tgtgcagcct ggaggatggg ttgttgggct ccccggcccg gctggcctcc cagctgctgg 780
gcgatgagct gcttctcgcc aaactgcccc ccagccggga aagtgccttc cgcagcctgg 840
gcccactgga ggcccaggac tcactctaca actcgccccc cacagagtcc tgcctttccc 900
ccgcggagga ggagccagcc ccctgcaagg actgccagcc actctgccca ccactaacgg 960
gcagctggga acggcagcgg caagcctctg acctggcctc ttctggggtg gtgtccttag 1020
atgaggatga ggcagagcca gaggaacagt gacccacatc atgcctggca gtggcatgca 1080
tcccccggct gctgccaggg gcagagcctc tgtgcccaag tgtgggctca aggctcccag 1140
cagagctcca cagcctagag ggctcctggg agcgctcgct tctccgttgt gtgttttgca 1200
tgaaagtgtt tggagaggag gcaggggctg ggctgggggc gcatgtcctg cccccactcc 1260
cggggcttgc cgggggttgc ccggggcctc tggggcatgg ctacagctgt ggcagacagt 1320
gatgttcatg ttcttaaaat gccacacaca catttcctcc tcggataatg tgaaccacta 1380
agggggttgt gactgggctg tgtgatggtg gggtgggagg gggcccagca acccccccacc 1440
ctccccatgc ctctctcttc tctgcttttc ttctcacttc cgagtccatg tgcagtgctt 1500
gatagaatca accccacctg gaggggctgg ctcctgccct cccggagcct atgggttgag 1560
ccgtccctca agggccctgc cagctgg 1587

<210> 50
<211> 669
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3580257CB1

<400> 50
cttgcctgaa caacaaacca actcaccact cctgacacca tgagtcacta cggcagctac 60
tacggaggcc tgggctacag ctgtggaggc ttcggtggcc tgggctatgg ctatggctgt 120
ggatgtggca gcttctgcag acggggttct ggctgtggct atggaggcta cggatatggc 180
tctggctttg gaagctacgg atatggctct ggctttggag gctacggata tggctctggc 240
tttggaggct atggatatgg ctgctgccgc ccatcgtaca atggaggata cggattctct 300
ggcttttatt aaatgaattg ctgaaattgg aagcagagga gaaacctcca aatgtgtttg 360
gtcctgtccc gtgctttcat tccaaaaatc cattctattg ccttcagcat caatggagag 420
atatttagct atgttaaatc tttaaaatag atttaagctg cttctgtgaa tatttgttgt 480
ctttttactt tcggaatccg tcacctgaaa tcatattcat gtataactac tctaggttgt 540

tcttctaaat ttgctttatt tttcttacca ccattacctt gatgttatct atcaagatcc 600
tagcaaaaac ttgtattttg cttttatcta ataaatgaag taaagtataa atattttaaa 660
aaaaaaaaa 669

<210> 51
<211> 1463
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3634758CB1

<400> 51
atcgcatctc agctggttgg ctttggttag agctcccgtc agactttcgt tcggccctag 60
gatttggtag ccccgaagtg tgggctctct ccagtaccag actcatttca gtaccagcct 120
ttgggaagtc gtgtgaatac ctcggtctct tagccacagg gatagaatgg cggcctgacg 180
gagccgcggc gccggcgaag tcgctgaggc gcgagctgga acccccagac cagctcaaac 240
gggagccaaa actcgaagct tggaagaatt agcaggaaat ggcggatgag gcgttgtttt 300
tgcttctcca taacgagatg gtgtctggag tgtacaagtc cgcggagcag ggggaggtgg 360
aaaacggacg atgtattact aagctggaaa acatggggtt tcgagtggga caaggattga 420
tagaaaggtt tacaaaagat actgcaaggt tcaaggatga gttagatatc atgaagttca 480
tttgtaaaga tttttggact acggtattca agaaacaaat cgacaatcta aggacaaatc 540
atcagggcat ctatgtactt caggacaaca aatttcgcct gcttactcag atgtctgcag 600
gaaaacagta tttagaacat gcatctaagt atttagcatt tacgtgtggc ttaatcagag 660
gtggcttatc aaacttggga ataaaaagta ttgtaacagc tgaagtgtct tcaatgcctg 720
cttgcaaatt tcaggtgatg atacagaagc tgtagaacat actgaaatgc aaggcttcaa 780
cagtgtaaag agataaatta ttcatgtaaa agtatttcaa gtagtgatga tttaattaca 840
ttgttcgatg tttgtacagg agtaagcatg tatttttatc aatttaacac agatcaaagg 900
agatgaaggg acattctgcc atgacataca cttaaccaaa actattcaaa atgaaaaccg 960
gatttcaaat aaccagacac caagatgcag ggcccttatt ttaaaccttt ttatttggtt 1020
agagtgatat gtatttagcc atagatggag aaacaaagct cagggtttgt tgaattagca 1080
tgagagaaaa ttatgtacca acagaattat ttgtgagaag aatgaacaaa ttttgataaa 1140
gtatgaattt gttttatttt aaaaagcaaa catactaaat tttttttatt ttattgctta 1200
taatttatta agaatgttta cacctgtata aggatttcat atatacattg tatgtgtgta 1260
tatataaata catatatgac tgcctaaatt gtttataaat ttaatttttc tttaataggt 1320
ttcattcctt cagagctcca ttaatgtaat caaaatgaaa tacagattag tttaaatgtg 1380
aattcagtga ctctagggcc aaagaatatt aggtatgttt ggaaagaatt tttgtattta 1440
ttcctgttac agttttgact ttc 1463

<210> 52
<211> 1686
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 4027923CB1

<400> 52
gtaacccatc gccacttggg atccaggtag taggtgcacg gccaggtgtg gatgcagaca 60
tccccgctca tctgttcctg ggtctgtgtg cgccaaggtc caggcctctt ctgcctctga 120
gggacgccag acgagtccca gagccgaccg cggcgcctcg gccgctcggc gcctgcgcag 180
tgaagaccgc ggcgagccgc gcgcatgcgt gcaggcccgg agccccaggc gctggtgggg 240
cagaaacgcg gcgccctgcg tcttctggtt ccgaggctgg tcctcaccgt ttccgctccg 300
gcggaagtga ggaggagggt ccttcgaccc gtgctgagct ggatggaccg cgagacgcgc 360
gccctcgccg acagccactt ccgaggcctg ggggtcgatg tccccggcgt cggccaggct 420
ccgggccggg tagccttcgt ctcggagccg ggcgccttct cctacgccga ctttgtgcgg 480
ggcttcttgc tgcccaacct gccctgcgtg ttttccagcg ccttcacgca gggctggggc 540
agccggcggc gctgggtgac gcccgcgggg aggcccgact tcgaccacct gctacggacc 600
tacggagacg tggttgtacc agttgcaaac tgtggggtcc aggaatacaa ctcgaacccc 660
aaagagcaca tgactctcag agactacatc acctactgga aagagtacat acaggcgggc 720
tactcctctc ccaggggctg tctctacctc aaagactggc acttgtgcag ggactttccg 780

gtggaggacg ttttcaccct gcctgtgtac ttctcgtccg actggctgaa tgagttctgg 840
gatgcactgg atgtggatga ctaccgcttt gtctacgcgg ggcctgcggg cagctggtcc 900
ccgttccatg ctgacatctt ccgctccttc agctggtctg tcaatgtctg tgggaggaag 960
aagtggctcc tcttcccccc agggcaggaa gaggccctgc gggaccgcca cggcaacctg 1020
ccctacgacg tgacctcccc agcactctgc gacacacacc tgcacccacg gaaccagctt 1080
gctggcccac ccttggagat cacgcaggaa gcgggcgaga tggtgtttgt gcccagtggc 1140
tggcaccacc aggtgcacaa cctggatgac accatctcca tcaaccacaa ctgggtcaat 1200
ggcttcaacc tggccaacat gtggcgcttc ttgcagcagg agctatgcgc cgtgcaggag 1260
gaggtcagcg agtggaggga ctccatgccc gactggcacc accactgcca ggtcatcatg 1320
aggtcctgct cgggcatcaa ctttgaagag ttttaccact tcctcaaggt catcgctgag 1380
aagaggctcc tggtcctgag ggaggcagcc gctgaggacg gtgctgggtt gggtttcgaa 1440
caggcagcct ttgatgttgg gcgcatcaca gaggtgctgg cctccttggt tgcgcacccc 1500
gacttccaga gagtggacac cagcgcgttc tcaccacagc ccaaagagct gctgcagcag 1560
ctgagagagg ctgttgatgc tgctgcggcc ccatagcacc tgtcgtgagg atagaaggac 1620
gggtggacga gaggcagcct cctgctccgg ggcccttcca gaaataaaga ccgccctccc 1680
tgtgaa 1686

<210> 53
<211> 2497
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 4348533CB1

<400> 53
gggcggcggg agctgctttg cctccaccga tctccctgtg cggccctcat gtgctgtgct 60
cgctgacacc cgaagtccgc ggctttccgc acacggtggg gtcgtcagac ccgctgccct 120
tggcggtcga agtcgtcgtg cgggcccgcg gcggccgccc atggagaagg ccaggagagg 180
cggggatggc gtccccccggg ggcccgtact gcacatcgtg gtggtcggat ttcaccacaa 240
gaagggctgc caggttgaat tctcttaccc gcccctgatt ccaggagatg gacatgacag 300
ccacacttta cctgaagaat ggaagtattt gcccttcctt gccttaccag atggcgcaca 360
caactaccag gaagatactg tgtttttca cttgccaccc agaaatggaa atggagccac 420
agtatttggt atctcttgct atcgacaaat tgaagccaag gcactgaaag taaggcaagc 480
agatatcacc agagagactg ttcagaaaag tgtctgtgtt ctaagcaagc tgcctctgta 540
tggtttactt caagcaaaac ttcaactcat tacacatgca tattttgaag agaaggattt 600
ttcccaaatt tctattctaa aggagcttta tgaacatatg aatagttcct tgggaggtgc 660
ttcattagaa ggatcccaag tatatcttgg tctgtcacct cgagatcttg tccttcattt 720
tcgacacaag gtcttaatcc tatttaagct aattcttctt gaaaaaaagg ttctttttta 780
tatttctcca gtgaataaat tggtgggtgc actgatgact gtgttatccc tttttccagg 840
catgattgaa catggtctca gtgactgttc tcagtataga ccccggaaaa gtatgtctga 900
agatggtggg cttcaggaaa gtaacccatg tgcagatgat tttgtttctg catccactgc 960
tgatgtttca cataccaact tgggaactat caggaaagtc atggcaggaa accatggaga 1020
agatgctgcc atgaagactg aggagccttt gttccaagtg gaagacagca gcaaagggca 1080
ggaacccaat gataccaatc aatatttgaa acctccatct cgcccatctc cagattcttc 1140
agaaagtgac tgggaaactt tggatcctag tgtcttagag gaccccaact tgaaagaaag 1200
ggaacagctg ggatcagacc agacaaattt gtttccaaag gactctgtcc cctcagagag 1260
tcttccaatt actgtacaac ctcaagctaa tacgggacag gtagtcctga taccagggct 1320
catttcgggt ttggaagagg atcagtatgg catgcccctg gccatcttca caaagggata 1380
tctgtgtttg ccttacatgg cattgcagca gcatcatctt ctctccgatg tcaccgttcg 1440
ggggtttgtt gctggagcta ctaacatcct ttttcgacaa cagaaacacc tcagtgatgc 1500
cattgtggaa gtagaagaag ctctgatcca gatccatgat ccagaactca ggaagctgct 1560
taacccaacc actgcagacc taaggttcgc agactaccta gtgaggcacg tgactgagaa 1620
tcgggatgac gtcttcctag atggcacggg ctgggaggga ggtgacgaat ggatccgggc 1680
ccagtttgcg gtctacattc atgccctgct ggctgccacg ctgcaattag acaatgaaaa 1740
gatattatcg gactatggga caacttttgt tacagcatgg aagaatactc acaactacag 1800
ggtgtggaac agcaacaagc atccagcact tgcagaaata aatccaaacc atccatttca 1860
aggccaatac tcagtatcag acatgaagtt aaggttctca cattctgttc agaatagtga 1920
acgtggcaaa aaaattggaa acgtcatggt cacaactagc cggaatgttg tacaaacagg 1980
aaaagctgtt ggccagtcag ttggaggagc tttttccagt gcaaagacag ctatgtcttc 2040
atggctttcc actttcacca cttccacctc ccaaagtctc actgagccac cagatgagaa 2100
gccttgagca aggcgtcaga ggctgctatt gctttctgag gtttaagtgt cccctgtctg 2160

tctgctgctc ccaggctgtt actagccaca gatccacagc aggggaccat atgtcgaact 2220
gtttacatgg atgttgctct aagtgaatgt ttcgggatgc cgaaatgatg aaatcacagc 2280
catagcaggg atggctttcc aggttggggt ttcaattgac tacttttatt tcagtctgag 2340
cctgattaaa acatacagtg aaccttctaa tgaattttga gttttggcat ttgaagtttc 2400
gtacattggt ttactttaga agagttttta cttatgtaaa ttttgttctg ttttggtgtt 2460
tgaatatcta gatggtcact gtaatttata cttgcta 2497

<210> 54
<211> 1783
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 4521857CB1

<400> 54
ccgctgtgct ctcgggaaag ccctgagaaa ctaccagaac ggggtccctg gggccagtgt 60
tgggagcgta cggtgaagtg gcagggctgc tgcgcccgtg cggaactccg ctgtggaaga 120
ttagcccctt tggtgactcg tcggttctgg ccgctttggg accctctact gtgcagtctg 180
aactttcagc aggtggagga gggcagagag agagctgaga tcgttggtcc atgccagtgc 240
gagagccagt acggccggaa gccatgatgg gactttcagt tcctttcttg gtggaggcca 300
gcacagggga gagggggtgc ccacagatgc tagaagctta aaggctgctg ggggctagcg 360
gcctcaggac ataaggtttc cagacaggac ggggcaggtg gtgatgcagc tgtagctcgt 420
gagtctgagc gggaaacagc ccccgcgcag tcggtgtcgc tgtgaacgga aaagggcctt 480
ccttacacag gaaacggggt ctctgggggа tccagtgacg tgcccaacca cagagaccag 540
acccctcctc ctgtagagag tggtgctgcc ctctcgggat gtacctgcag gtggagaccc 600
gcaccagctc ccgcctccat ctgaagaggg ctccaggcat ccggtcctgg tccctgctgg 660
ttggaatctt gtcgattggc ctggctgctg cctactacag cggagatagc ctgggctgga 720
agctcttcta cgtcacaggc tgcctgtttg tggctgtgca gaacttggag gactgggagg 780
aagccatctt cgacaagagc acagggaagg ttgttttgaa gacgttcagc ctctacaaga 840
agctgctgac tcttttcaga gctggccacg accaggtggt ggtcctgctc catgttgtcc 900
ccgacacagc gtcctctccc tggtggacat ccccagcggt caggtgcttc cccaagggca 960
gtgaggggtg aacatccagg gcctacctgg ctgtgcacgc tgcagccaca ctgtggaagc 1020
tgcccctccc cgaggacccg cctcccttgc tgatgccagg atctcggcgc atagaccact 1080
ctgccccagc ggtcgtcaca gaaaggtctc tctgttcctc acactcagct tcagcataag 1140
ctgtgaggcc agaaaaaagg tcagctcttc tagtatcgtg cagtgcttaa aaaccgggag 1200
ctccagccgg gcgcagtggt tcatgccagt aatcccagca ctttcggagg ccgaggtggg 1260
aggattgctt gaggccagga gttcaagacc agcctgggca acacagcaag atcctgtctt 1320
tgtaaaaaaa ctaaccaaac aggaaaaact gggagatttt ctgcagaaat tgagttccag 1380
cctctctcga acctgggaag acctggcagg agggggctgg gctctcggct acagacttct 1440
ccccaccccg taggagctga acgccaacca tcctgacccg ccagtgctct tggtctcctg 1500
agtgtaccca ggtcctccca ggtgcggtgt gcaccgagcg cgcctggcct gatgccctgg 1560
cctgtgagct ggggactcct gggccctgtg agcccctagg cggcaggccc aggaatggct 1620
gggtaggaca gggaacacct ttgccccacg tctggctgtg acctcggtga aagccgacag 1680
gagagagatg ggaccctcct cctcagtagt ggctgccagt ccctcgttgc aggacagggt 1740
catcataacc ataaataacc cttcacgtgt caaaaaaaaa aaa 1783

<210> 55
<211> 1461
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 4722253CB1

<400> 55
ccagactcgg tcccgcgggc tttaaggggc ccgggcggcg aaggcgcacg gagccaagtt 60
ccgggcagca gccctgggat gccctgagcc gggccgggcg gaggtctccg ttgcctagca 120
accggggccg cggcctgtgg gcggagcctg caccgtggct gcgacacggg gcggggcctc 180
agcgggagcc ggccgaggag cgggcaccgg ccattggcac aggcacgggc cattggcgca 240
tgcgtagggc gcggccttgg ggcggggccc acccagtggc ggagtcttct gaggggcggg 300

tcgtgggcgg ggcctggagg cggagccggc gccgtcagta gtcggctgag gacggaggcg 360
gggctgggcc tgggatgggg cggggcctgc gctggtgggg cgggcggggg aggagacacg 420
gccaagcgcc cgagtggggg ccgttggttg gtgcgcggct gaagggtgtg gcgcgagcag 480
cgtcgttggt tggccggcgg cgggccggga cgggcatggc cctgctgctg tgcctggtgt 540
gcctgacggc ggcgctggcc cacggctgtc tgcactgcca cagcaacttc tccaagaagt 600
tctccttcta ccgccaccat gtgaacttca agtcctggtg ggtgggcgac atccccgtgt 660
caggggcgct gctcaccgac tggagcgacg acacgatgaa ggagctgcac ctggccatcc 720
ccgccaagat cacccgggag aagctggacc aagtggcgac agcagtgtac cagatgatgg 780
atcagctgta ccaggggaag atgtacttcc ccgggtattt ccccaacgag ctgcgaaaca 840
tcttccggga gcaggtgcac ctcatccaga acgccatcat cgaaagccgc atcgactgtc 900
agcaccgctg tggcatcttc cagtacgaga ccatctcctg caacaactgc acagactcgc 960
acgtcgcctg ctttggctat aactgcgagt cctcggcgca gtggaagtca gctgtccagg 1020
gcctcctgaa ctacataaat aactggcaca aacaggacac gagcatgaga ccacgctcct 1080
ctgccttctc ctggcctggg acacacagag ccaccccggc cttcctggta tcgccagcct 1140
taaggtgtct ggagccccca cacttggcca acctgacctt ggaagatgct gctgagtgtc 1200
tcaagcagca ctgacagcag ctgggcctgc cccagggcaa cgtgggggcg gagactcagc 1260
tggacagccc ctgcctgtca ctctggagct gggctgctgc tgcctcagga ccccctctcc 1320
gaccccggac agagctgagc tggccagggc caggagggcg ggagggaggg aatgggggtg 1380
ggctgtgcgc agcatcagcg cctgggcagg tccgcagagc tgcgggatgt gattaaagtc 1440
cctgatgttt ctcaaaaaaa a 1461

<210> 56
<211> 2116
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 4878134CB1

<400> 56
cgcccgcgga gtcgctgagg aggcggaaga ctgggtactc ggatccggag cctgagtcgc 60
cgcccgcgcc ggggcgtggc cccgcaggct ctccggccca tctccacacg ggcaccttct 120
ggctgacccg gatcgtgctc ctgaaggccc tagccttcgt gtacttcgtg gcattcctgg 180
tggctttcca tcagaacaag cagctcatcg gtgacagggg gctgcttccc tgcagagtgt 240
tcctgaagaa cttccagcag tacttccagg acaggacgag ctgggaagtc ttcagctaca 300
tgcccaccat cctctggctg atggactggt cagacatgaa ctccaacctg gacttgctgg 360
ctcttctcgg actgggcatc tcgtctttcg tactgatcac gggctgcgcc aacatgcttc 420
tcatggctgc cctgtggggc ctctacatgt ccctggttaa tgtgggccat gtctggtact 480
ctttcggatg ggagtcccag cttctggaga cggggttcct ggggatcttc ctgtgccctc 540
tgtggacgct gtcaaggctg ccccagcata cccccacatc ccggattgtc ctgtggggct 600
tccggtggct gatcttcagg atcatgcttg gagcaggcct gatcaagatc cgggggacc 660
ggtgctggcg agacctcacc tgcatggact tccactatga gacccagccg atgcccaatc 720
ctgtggcgta ctacctgcac cactcaccct ggtggttcca tcgcttcgag acgctcagca 780
accacttcat cgagctcctg gtgcccttct tcctcttcct cggccggcgg gcgtgcatca 840
tccacggggt gctgcagatc ctgttccagg ccgtcctcat cgtcagcggg aacctcagct 900
tcctgaactg gctgactatg gtgcccagcc tggcctgctt tgatgacgcc accctgggat 960
tcttgttccc ctctgggcca ggcagcctga aggaccgagt tctgcagatg cagagggaca 1020
tccgaggggc ccggcccgag cccagattcg gctccgtggt gcggcgtgca gccaacgtct 1080
cgctgggcgt cctgctggcc tggctcagcg tgcccgtggt cctcaacttg ctgagctcca 1140
ggcaggtcat gaacacccac ttcaactctc ttcacatcgt caacacttac ggggccttcg 1200
gaagcatcac caaggagcgg gcggaggtga tcctgcaggg cacagccagc tccaacgcca 1260
gcgcccccga tgccatgtgg gaggactacg agttcaagtg caagccaggt gaccccagca 1320
gacggccctg cctcatctcc ccgtaccact accgcctgga ctggctgatg tggttcgcgg 1380
ccttccagac ctacgagcac aacgactgga tcatccacct ggctggcaag ctcctggcca 1440
gcgacgccga ggccttgtcc ctgctggcac acaacccctt cgcgggcagg cccccgccca 1500
ggtgggtccg aggagagcac tacaggtaca agttcagccg tcctgggggc aggcacgccg 1560
ccgagggcaa gtggtgggtg cggaagagga tcggagccta cttccctccg ctcagcctgg 1620
aggagctgag gccctacttc agggaccgtg ggtggcctct gcccgggccc ctctagacgt 1680
gcaccagaaa taaaggcgaa gacccagccc ctcggcggct cagcaacgtt tgcccttccc 1740
tgcgcccagc ccaagctggg catcgccaag agagacgtgg agaggagagc ggtgggaccc 1800
agcccccagc acgggggtcc agggtggggt ctgttgtcac atactgtggc ggctcccagg 1860
ccctgcccac ctggggcccc acatccaggc caacccttgt gcccaggcgc caggggctct 1920

gatctcccat ccatcccacc ctcctcccag aggcccagcc tggggctgtg ccgcccacag 1980
gagttgagac aatggccatc ctgacacctt cctccactac agccctgacc atagacccag 2040
ccaggtagct cttggggtct ctagcgtccc agggcctggt ttctgttccc tcttcaatgg 2100
tgtgttccca gccagg 2116

<210> 57
<211> 702
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 5050133CB1

<400> 57
ccgaggcccg ggcgcaacca cgggctccca ggcagcctcc gccagccgga ccccgtcgcc 60
ctcctgatgc tgctcgtgga cgctgatcag ccggagccca tgcgcagcgg ggcgcgcgag 120
ctcgcgctct tcctgacccc cgagcctggg gccgaggcga aggaggtgga ggagaccatc 180
gagggcatgc tcctcaggct ggaagagttt tgcagcctgg ctgacctgat caggagtgat 240
acttcacaga tcctggagga aaacatccca gtccttaagg ccaaactgac agaaatgcgt 300
ggcatctatg ccaaagtgga ccggctagag gccttcgtca agatggttgg acaccacgtc 360
gccttcctgg aagcagacgt gcttcaggct gagcgggacc atggggcctt ccctcaggcc 420
ctgcggaggt ggctgggatc cgcagggctc ccctccttca ggaacaagtc acctgcaccg 480
gtgcccgtga cgtacgagct gcccacactg tataggacgg aggactattt tcctgtggac 540
gccggggaag cacagcacca cccccgcacc tgccctcggc ctttgtgagc tttgtggtct 600
tcccatcagg aacgctggaa agtgacattg tgtacacgct gcagcttggg ggtttttttct 660
ttgtattgct gtttatttta tattttaaaa atatttaaaa aa 702

<210> 58
<211> 2613
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 5630124CB1

<400> 58
tctcggaaca cattttactg ggctcgagtt tagccgccac cggagggtgg gggaccttga 60
gtcatgctcc tataaccacc gctagagttc ctcgtctttg agtgcagagg tttagactgt 120
gtctttgtgt gcagaaagtc ctgcagttct cacagcgacc tgccagaaaa agtcgttccc 180
aaatgtttgt aaatcctccg ttgggcaacc cgccttcacg ttctgcggtg atcttgtcga 240
gcgactaagc gtgcagtatt agcagagaag ggggtggcag agtgctggcg ctgaaggtca 300
tgttgcatgg gtaactgtcg tgttgtaggg gcggggaaga gggaggagac actgaccacc 360
ccagaggccg ccccattagc tcgcttgctt tgggcggcgt cgctcccacg gcgcccaggg 420
tacccccgcc gctgtctgcc tgtcttcctc cattaccgcg caggcttggt caccgcatta 480
aggcattccc gctctccgcg gaactgctct gccgtctcgg cggtgaaagt gtgagagggt 540
ccgtagttgg gtcaactttg actcctctcg cctgcccgga tccttaaggg cctcctcgtc 600
ctcccggtct ccggtcgctg ccgggtctgt gcgccggtcc gcgcccgccc tcgctctgcc 660
atgggcgctt ccagctcctc cgcgctggcc cgcctcggcc tcccagcccg gccctggccc 720
aggtggctcg gggtcgccgc gctaggactg gccgccgtgg ccctggggac tgtcgcctgg 780
cgccgcgcat ggcccaggcg gcgccggcgg ctgcagcagg tgggcaccgt ggcgaagctc 840
tggatctacc cggtgaaatc ctgcaaaggg gtgccggtga gcgaggctga gtgcacggcc 900
atggggctgc gcagcggcaa cctgcgggac aggttttggc tggtgattaa ggaagatgga 960
cacatggtca ctgcccgaca ggagcctcgc ctcgtgctca tctccatcat ttatgagaat 1020
aactgcctga tcttcagggc tccagacatg gaccagctgg ttttgcctag caagcagcct 1080
tcctcaaaca aactccacaa ctgcaggata tttggccttg acattaaagg cagagactgt 1140
ggcaatgagg cagctaagtg gttcaccaac ttcttgaaaa ctgaagcgta tagattggtt 1200
caatttgaga caaacatgaa gggaagaaca tcaagaaaac ttctccccac tcttgatcag 1260
aatttccagg tggcctaccc agactactgc ccgctcctga tcatgacaga tgcctccctg 1320
gtagatttga ataccaggat ggagaagaaa atgaaaatgg agaatttcag gccaaatatt 1380
gtggtgaccg gctgtgatgc ttttgaggag gatacctggg atgaactcct aattggtagt 1440
gtagaagtga aaaaggtaat ggcatgcccc aggtgtattt tgacaacggt ggacccagac 1500

actggagtca tagacaggaa acagccactg gacaccctga agagctaccg cctgtgtgat 1560
ccttctgaga gggaattgta caagttgtct ccacttttg ggatctatta ttcagtggaa 1620
aaaattggaa gcctgagagt tggtgaccct gtgtatcgga tggtgtagtg atgagtgatg 1680
gatccactag ggtgatatgg taaagggctt cagcaaccag gagggattga ctgagatctt 1740
aacaacagca gcaacgatac atcagcaaat ccttattatc cagccttcaa ctatctttac 1800
cctggaaaac aatctcgatt tttgactttt caaagttgtg tatgctccag gttaatgcaa 1860
ggaaagtatt agaggggga atatgaaagt atatatataa attttaggta ctgaaggctt 1920
taaaaataat taagatcatc aaaaatgcta ttttgaatgt tatcatggct attacacttt 1980
tacttcctga ctttaatatt gatgaataaa gcaagtttaa tgaatcaact aaaaagctgc 2040
aaaaatgttt ttaaaatgtg tgccttttat tacctatcag tctatgtttt gggagaaatg 2100
ggaagcaaca gatcactgtg tcctgatgtg caggacgcat gttaccacac tcacaaatgc 2160
ctaatattgg tctttatgtg gccattgagt cctgttgact ttccactcat gtgcttttta 2220
ctctagcatt atggaatctg ggctgtactt gagtatggaa attctcttat agacttagtt 2280
ttagtactct attacacctt tactaagcca ataaaagta atctgtttgt gtgtaactgc 2340
cagatatacc acctggaatt ccaagtaaga taaggaagag gatgacattt aaaagagaat 2400
ggaattttga gagtaggaat gcaaggaaga cagcatgaac atatttttt cagtgcaaat 2460
aatttttcg taacaaagaa acgaacaact ttggtatgat cttaagcaaa aatactcact 2520
gaaatagtat gtggatgaat tcacctactt acaattttat ggtttctttg taaataataa 2580
atgtgaatct caatcctgaa aaaaaaaaaa aaa 2613

<210> 59
<211> 1778
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 5677286CB1

<400> 59
gtctgccctt tgaatggtgt gatctgggct caccgcaacc tccgcctcca gggttcaagc 60
gattctcctg cctcagcctc cgagtagctg ggattacagg catgcgccac cacgcccgga 120
taattttgta tttttagtag agacagggtt tctccatgtt ggtcaggctg gtctcaaact 180
gacctcaggt gatccgccca cctcggcctc ccaaagtgct gggattacag gcatgagcca 240
ctgcgcctgg ccggacaata atgttttgag ctgttgctga gtctggcttt tgtgacggga 300
acttttagca ctcaaggagt ggccttgtct ctggcctgtc cccgagtgct catgggcagc 360
caaggcctgg ggactgctca ggcagggtag atgtatttgc caggccagct cccgtccctg 420
ggacctcagg gagtagcttc atccctgaag gctgagtgtc ttctgcctca tgtgggtggc 480
ccctctgcag ggttctcagg catccaggga agtgggtgcc acagcgttgc ctccagcctc 540
actggggctg gctggacttt gctctgacaa cacagcagtg gctgggcccc tgcaactctc 600
cagggctcag tttccctgtt ggccgcacta tggcataggc ccctgtgtgg acaccgatga 660
gctgacccca caaatgccac ccggccgctc ccccaggctt cagtggccta acagctgttg 720
tgtcaggaac cagcttaaag aatctttctt gctttctcaa actctccagg aattcctggt 780
gactggaagg gggagtgact caggccctca acctctagca ggtaagggtg ctacctcctg 840
ggtgggcatg cccagttcct cagtgggccc agcgcggccc agcctccgag gaagctctga 900
gcagctggct tggctccgag gtatttttag ctacagatcc agcccccctc catcaccaca 960
acagtcgagt ccaaatcaaa cgcgctccag gcaggtgggg ctggggttct gccagcctcc 1020
tggccagcag gggtgggtgg gcagactggg gccagtatca gctgttctgc ctggacccag 1080
gccgggctgg gaaggcacac ttgtgcttat ttcccgcctc cacttctgtg caagcttgtg 1140
ctgtcataag cagagatcac agccccattt cttggatgga gaaagtggac actgaggtct 1200
gaggcttttg aggacagtca ggagccctcc tatgggctcc agtgatgcac tcaccagctt 1260
ctggtcctct tcttccacct cttagagtgc cttggctccc tcctgtcgtc ctggggaacc 1320
tcggccccag ccctgcctcc ccagccagtc acagctcctc cctggtcacc ctgagggagc 1380
tcagggcccg gctggtagct gggttgctct gcttctgtcc ccgactcctg tggagcctgg 1440
caggcaactc catgatctga ccccggttac cttgacagcc ctgcctggcc tcccctctca 1500
tggcccagcc accccagaac ctgaagaggt tttctagctg ccatgcattt gccaggctgg 1560
gttacccacc ctactttccc tgcctgccct ccagtgctgc caggcctagt gtgccagcca 1620
gcgctcagcc ttcagtaaag ggttcccctg cttccaacct ccattgcact gcttccccta 1680
agactgtgac ctcctggaag gctggagcac aactgcctct caataaacgt gttgcaaaaa 1740
aggaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaa 1778

<210> 60
<211> 1234

<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 6436791CB1

<400> 60
cctgcagacg acggcgtcgt gggtggtcac cgttatccct taggtctgga gaggggacat 60
ccgagcgagg gccacttgcg gccaggcccg agctcgtcca gctccgggtg accacagagt 120
gccgcgggcg ggcagagggg ccggaaaccc aggccgcttc gtccctgttt ccggcagcgc 180
cgcgctgctc cgggagccgc tgtggcagcg tatgctgcca cggggactga agatggcgcc 240
gcgaggtaaa cggttgtcct ccaccccgct ggaaatcctg ttctttctga acgggtggta 300
taatgctacc tatttcctgc tggaactttt catatttctg tataaaggtg tcctgctacc 360
atatccaaca gctaacctag tactggatgt ggtgatgctc ctcctttatc ttggaattga 420
agtaattcgc ctgtttttg gtacaaaggg aaacctctgc cagcgaaaga tgccactcag 480
tattagcgtg gccttgacct tcccatctgc catgatggcc tcctattacc tgctgctgca 540
gacctacgta ctccgcctgg aagccatcat gaatggcatc ttgctcttct tctgtggctc 600
agagctttta cttgaggtgc tcaccttggc tgctttctcc agtatggaca cgatttgaag 660
tacagaattt cagccagcag cccatcaggc tgacaccaca catattgctt ctggtacttt 720
agccacacca gtgagaattg gtggggcaag ttgtcctgag aaaggctgtg tggcttttct 780
tcagcacaga catttgggca agcaactcag cataaggcca gtgggtacca tcttctaaac 840
caggaccatc agcccaagag actcttctac actccagtat agggaggggc aaggttattc 900
ccatcctgcc ccttctcaga accagtcccc tgctgacctc aagttctcct ccttgatcac 960
cgtggccaga gcatctcgtg tggaccatct aggctccttg ggcttcaagc aggacctgag 1020
ccacatgctc cctgtacgag ctgtgctata cctgtcccac atgagcacgg agagcctcat 1080
gttggtgggt ttccagagtg atgtgaaagc ctctcacccc aatcctcgga gactgagttc 1140
cacaactttt ttagtagctc atagtgttat ttttctactc tcttcatgaa actaacttta 1200
ttttataata aatatatatt ttctgttaaa aaaa 1234

<210> 61
<211> 863
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1820972CB1

<400> 61
ggaggtgcct cagccatggc atggatccct ctcttcctcg gcgtccttgc ttactgcaca 60
ggatccatgg actcctttga attgactcag gcaccgtcaa cgtccgtgtc cccaggacag 120
acagccacca tctcctgctc tggcgagaag gtgggaagta aattcttttc gtggtatcaa 180
cagaaggaag gccagtcccc tgtcgtaatc atctatcaga atgggaagcg gccctcagag 240
attgctgacc gattctctgg ctccaagtct ggggacacgg ccaccctgac catcagcaga 300
gcccaagccg gggatgaggc tgactatttc tgtcaggtgt gggacagcag cactgcggtg 360
ttcggcggag ggaccaagct gaccgtccta ggtcagccca aggctgcccc ctcggtcact 420
ctgttcccgc cctcctctga ggagcttcaa gccaacaagg ccacactggt gtgtctcata 480
agtgacttct acccgggagc cgtgacagtg gcctggaagg cagatagcag ccccgtcaag 540
gcgggagtgg agaccaccac accctccaaa caatgcaaca acaagtacgc ggccagcagc 600
tacctgagcc tgacgcctga gcagtggaag tcccacagaa gctacagctg ccaggtcacg 660
catgaaggga gcaccgtgga gaagacagtg gcccctacag aatgttcata ggttctaaac 720
cctcaccccc cccacgggag actagagctg caggatcccc aggggagggg tctctcctcc 780
caccccaagg catcaagccc ttctcccgtg cactcaataa accctcaata aatattctca 840
tttgtcactc aaaaaaaaaa aaa 863

<210> 62
<211> 2521
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature

<223> Incyte ID No: 3286805CB1

<400> 62
gcttgctggc cgctgtgtag ggctggtgag tggctggggc tgtctgagcc atgaacaact 60
tcagggccac catcctcttc tgggcagcgg cagcatgggc taaatcaggc aagccttcgg 120
gagagatgga cgaagttgga gttcaaaaat gcaagaatgc cttgaaacta cctgtcctgg 180
aagtcctacc tggaggggge tgggacaatc tgcggaatgt ggacatggga cgagttatgg 240
aattgactta ctccaactgc aggacaacag aggatggaca gtatatcatc cctgatgaaa 300
tcttcaccat tccccagaaa cagagcaacc tggagatgaa ctcagaaatc ctggaatcct 360
gggcaaatta ccagagtagc acctcctact ccatcaacac agaactctct ctttttttcca 420
aagtcaatgg caagttttcc actgagttcc agaggatgaa gaccctccaa gtgaaggacc 480
aagctataac tacccgagtt caggtaagaa acctcgtcta cacagtcaaa atcaacccaa 540
ctttagagct aagctcaggt tttaggaagg aactccttga catctctgac cgtctagaga 600
acaaccagac gaggatggcc acctacctgg cagaactcct ggtgctcaac tatggcaccc 660
acgtcaccac cagtgtcgac gctggggctg ctcttattca ggaggaccac ctcagggcct 720
ccttcctcca agacagccag agcagtcgta gtgccgtgac cgcctctgct ggacttgcct 780
ttcaaaacac cgtgaacttc aaatttgagg aaaactatac ctcgcagaat gtcctcacca 840
agagctacct ctcaaaccga accaactcca gggtgcagag cattggaggg gttccttttt 900
acccaggcat caccctccag gcctggcagc agggtatcac caaccacctg gtggccatcg 960
accgctctgg cctgccgctg catttcttca tcaaccccaa catgctacct gacttgccag 1020
gccccctggt gaagaaggtg tcaaagacag tggaaactgc tgtgaagcgc tattatacat 1080
tcaacaccta ccctggctgc acagatctca attctcccaa cttcaatttt caggccaaca 1140
cggatgatgg ctcctgcgag gggaaaatga ccaacttctc tttcggtggg gtttatcagg 1200
aatgcactca gctctcaggg aatagggatg tcctcctctg ccaaaagttg gagcagaaga 1260
atccactcac tggtgatttc tcctgcccct ctggctactc cccggtgcac ctgttatccc 1320
agatccacga ggagggttac aaccacctgg agtgtcatcg aaagtgcact ctcctcgtct 1380
tctgcaagac cgtgtgtgaa gatgtgttcc aggtggcaaa agctgaattt agggcttttt 1440
ggtgtgtggc cagcagccaa gtacctgaaa actcaggact gctttttggg ggcctcttca 1500
gcagcaagag cataaacccc atgacaaatg cacagtcatg cccagccggc tactttccac 1560
tgagactctt tgaaaacctc aaggtatgtg tttctcagga ctatgagttg ggaagcaggt 1620
ttgcggtccc ctttggcggg ttctttagct gcacagttgg gaaccccctg gtagatcctg 1680
ctatatccag agatttaggg gcaccgtctc tgaaaaagtg ccccgggggc ttcagccagc 1740
acccagccct catcagcgat ggatgccaag tgtcctattg cgtcaaatcc gggctcttca 1800
caggagggtc cctgcccccc gccaggctcc caccttcac ccggccaccc ctcatgagtc 1860
aggctgccac caatactgtc atagtgacca attctgagaa tgcgagatcc tggattaaag 1920
actcccagac ccaccagtgg aggctgggag aaccgataga gctgcggagg gccatgaatg 1980
tcatccatgg ggatggtggt ggtctgtcag gaggggctgc agctggggtc acagtggggg 2040
tcaccaccat tctggctgtt gttatcacct tggccatcta cggcacccgg aagttcaaga 2100
agaaagcata tcaggcaatt gaggaaaggc agagtttggt tccaggcact gcagcaactg 2160
gagacaccac ttaccaagag caggggcaga gtccagctta aatctctccc cgaaaatggt 2220
ttctctcatc tccagtgtgg tcattgctga ccactctgtt ttcctaagca ttgaaatggc 2280
aagtgcaacc aaaaatcttg ggctgagaag aggcccttca cccaaggagt ttgaagtatc 2340
acagcgtgtg ggaaggtggg aaccaggata cccattcatt tccaaccgag acacagagaa 2400
gtgagtcaca gaatttgagc ccgctctctt gactgcccag ccagagacac tgatttctgt 2460
aacctcttca cttgatcctg cctcttaagc attaaaacat tctcctaaaa aaaaaaaaaa 2520
a 2521

<210> 63
<211> 1765
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3506590CB1

<400> 63
gcactggaag tcgccggtgt ttccattcgg tgatcatcac tgaacacaga ggactcacca 60
tggagtttgg gctgagctgg gttttcctcg ttgctctttt aagaggtgtc cagtgtcagg 120
tgcagctggt ggagtctggg ggaggcgtgg tccagcctgg gaggtccctg agactctcct 180
gtgcagcctc tggattcacc ttcagtagct atgctatgca ctgggtccgc caggctccag 240
gcaaggggct ggagtgggtg gcagttatat catatgatgg aagcaataaa tactacgcag 300
actccgtgaa gggccgattc accatctcca gagacaattc caagaacacg ctgtatctgc 360

aaatgaacag cctgagagct gaggacacgg ctgtgtatta ctgtgcgaga gcgggagaag 420
ggagccccga tactttggtt gcttttgata tctggggcca agggacaatg gtcaccgtct 480
cttcagcttc caccaagggc ccatcggtct tccccctggc gccctgctcc aggagcacct 540
ctgggggcac agcggccctg ggctgcctgg tcaaggacta cttccccgaa ccggtgacgg 600
tgtcgtggaa ctcaggcgcc ctgaccagcg gcgtgcacac cttcccggct gtcctacagt 660
cctcaggact ctactccctc agcagcgtgg tgaccgtgcc ctccagcagc ttgggcaccc 720
agacctacac ctgcaacgtg aatcacaagc ccagcaacac caaggtggac aagagagttg 780
agctcaaaac cccacttggt gacacaactc acacatgccc acggtgccca gagcccaaat 840
cttgtgacac acctcccccg tgcccacggt gcccagagcc caaatcttgt gacacacctc 900
ccccatgccc acggtgccca gagcccaaat cttgtgacac acctcccccg tgcccaaggt 960
gcccagcacc tgaactcctg ggaggaccgt cagtcttcct cttcccccca aaacccaagg 1020
ataccttat gatttcccgg acccctgagg tcacgtgcgt ggtggtggac gtgagccacg 1080
aagaccccga ggtccagttc aagtggtacg tggacggcgt ggaggtgcat aatgccaaga 1140
caaagctgcg ggaggagcag tacaacagca cgttccgtgt ggtcagcgtc ctcaccgtcc 1200
tgcaccagga ctggctgaac ggcaaggagt acaagtgcaa ggtctccaac aaagccctcc 1260
cagcccccat cgagaaaacc atctccaaag ccaaaggaca gccccgagaa ccacaggtgt 1320
acaccctgcc cccatcccgg gaggagatga ccaagaacca ggtcagcctg acctgcctgg 1380
tcaaaggctt ctaccccagc gacatcgccg tggagtggga gagcaatggg cagccggaga 1440
acaactacaa caccacgcct cccatgctgg actccgacgg ctccttcttc ctctacagca 1500
agctcaccgt ggacaagagc aggtggcagc aggggaacat cttctcatgc tccgtgatgc 1560
atgaggctct gcacaaccgc tacacgcaga agagcctctc cctgtctccg ggtaaatgag 1620
tgccatggcc ggcaagcccc cgctccccgg gctctcgggg tcgcgcgagg atgcttggca 1680
cgtaccccgt gtacatactt cccaggcacc cagcatggaa ataaagcacc cagcgctgcc 1740
ctgggccccc gcaaaaaaaa aaaaa 1765

<210> 64
<211> 1264
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 003600CB1

<400> 64
ccctgctcca gtcacacccg gaagctgact ggtccacgca cagctgaagc atgaggaaac 60
tcatcgcggg actaattttc cttaaaattt agacttgcac agtaaggact tcaactgacc 120
ttcctcagac tgagaactgt ttccagtata tacatcaagt cactgagatc tccagcaccc 180
tgccggtggc actactgaga gacgaggtgc cagggtggtt cctgaaagtg cctgagcccc 240
aacttatcag caaggagctc atcatgctga cagaagtcat ggaggtctgg catggcttag 300
tgatcgcggt ggtgtccctc ttcctgcagg cctgcttcct caccgccatc aactacctgc 360
tcagcaggca catggcccac aagagtgaac agatactgaa agcggccagt ctccaggttc 420
ccaggcccag ccctggccac catcatccac ctgctgtcaa agagatgaag gagactcaga 480
cagagagaga catcccaatg tctgattccc tttacaggca tgacagcgac acaccctcag 540
atagcttgga tagctcctgc agttcgcctc ctgcctgcca ggccacagag gatgtggatt 600
acacacaagt cgtcttttct gaccctggag aactaaaaaa tgactccccg ctggactatg 660
agaacataaa ggaaatcaca gattatgtca atgtcaatcc agaaagacac aagcccagtt 720
tctggtattt tgtcaaccct gctctgtctg agccagcgga atatgatcaa gtggccatgt 780
gaattccaaa tatttttaat ggggtccagt tctctatgga ttcttacatt taatttgtag 840
ggaaatgcca tttttccccc ttaaacaagg catggggctc acaagtctat ggagacaggc 900
caaaaagaat gtggagaaga aaactgataa atacacagag gtcctcaaga cccatggact 960
cctggtctgt acccaaaaaa gctgttcgtt cctcaaaaac aaaaacaagg cttggctggg 1020
aaaacaggcc aatgccccgg caagaaaggt tgagatcaga tgttaggaag aactttcagg 1080
taaagtatga gaactatgga gtccatcagc agagatagta gtgaagtctc tccccaggga 1140
aaatttaaa aaggttgaat cagctgttgt agagttctat ttggcaatct catggttaaa 1200
tgacttccct ttgagctctt taattattgg caataaacaa cttctttaaa agttttaaat 1260
aaaa 1264

<210> 65
<211> 3415
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1251534CB1

<400> 65
gcctcctcca cctcctcgtc tttctcccgg gaaccttgac gacgccttcc gcttggccct 60
gccttctgcc gcatccccgc cgccgcggcg ccttgaggag caggagaaga cgcagccggg 120
ccgccgccgt tagaggggtt cccggccgcc gctcgccccg tcggccgcca ccgcctccgg 180
ggtcagccct ctctctgggt ctccgctttc tcctgccgcc agcgcccgct catcgccgcg 240
atggggctcc tggactcgga gccgggtagt gtcctaaacg tagtgtccac ggcactcaac 300
gacacggtag agttctaccg ctggacctgg tccatcgcag ataagcgtgt ggaaaattgg 360
cctctgatgc agtctccttg gcctacacta agtataagca ctctttatct cctgtttgtg 420
tggctgggtc caaaatggat gaaggaccga gaacctttтc agatgcgtct agtgctcatt 480
atctataatt ttgggatggt tttgcttaac ctctttatct tcagagagtt attcatggga 540
tcatataatg cgggatatag ctatatttgc cagagtgtgg attattctaa taatgttcat 600
gaagtcagga tagctgctgc tctgtggtgg tactttgtat ctaaaggagt tgagtatttg 660
gacacagtgt tttttattct gagaaagaaa aacaaccaag tttctttcct tcatgtgtat 720
catcactgta cgatgtttac cttgtggtgg attggaatta agtgggttgc aggaggacaa 780
gcatttttg gagcccagtt gaattccttt atccatgtga ttatgtactc atactatggg 840
ttaactgcat ttggcccatg gattcagaaa tatctttggt ggaaacgata cctgactatg 900
ttgcaactga ttcaattcca tgtgaccatt gggcacacgg cactgtctct ttacactgac 960
tgccccttcc ccaaatggat gcactgggct ctaattgcct atgcaatcag cttcatattt 1020
ctctttctta acttctacat tcggacatac aaagagccta agaaaccaaa agctggaaaa 1080
acagccatga atggtatttc agcaaatggt gtgagcaaat cagaaaaaca actcatgata 1140
gaaaatggaa aaaagcagaa aaatggaaaa gcaaaaggag attaaattga actgggcctt 1200
aactgttgtt gacagtgagg aaaaactccc atatcatata aaatttcagg gaaaacagaa 1260
gcaaaggaga gcttgggggt ggggagaaaa gacaaatgtg ctctatgtcc tagtaactct 1320
tagactgagt aaagtgttaa taccataccc agatgtttta tttatgaagt ttttatttta 1380
aacatttttt ttaaaaatta gccttgatat tctccagacc aaagcaatca ttaagtgact 1440
ttggggattc tccccctgtt cacatccagt tgtctaaagg atgagatttt tcatgtatct 1500
tatagtcact cattcttcgg tctgaatttt agacgatcac agaaacggtc tttatgaatt 1560
attttgataa attactaatt atcttatcta ctgactgaaa tcagtggtgt tacattttct 1620
tgtccaaagc tgaaaatgtg tatacactta aacttgcaca tttgaattca tttgctgacc 1680
ggaatggtca aatctctcca cctctagtca gagtataatt ttggttgtaa ttaaattttt 1740
aaaatctgct gatctctgta gaatcttaga ggcttgatga tgatggtgtt ggtgaaaata 1800
agaaagaatt gcagtaaagt cttgtctggt gacccagaga tcaccatgac ttgaggcaca 1860
aatcactgtg gggaaacaat tttttgtgat gaaaaggcag catttgaata ctcctgttag 1920
tagcagaaat atattatgaa aattaagatt attgtctgat tgaaacatga aacaactcat 1980
gtctttatta gtaacatcat aagatagtta catttatgtg ctgttagaat atgttgattt 2040
ttatcaggct ttccttgttt tgatttatgg ctgttcctga tttttcatat gtggaaatat 2100
acctacctct tccgttggaa agaacattta aaattaaata aattttaatt aaaaaatcaa 2160
ggagtcttct aatgtaaatt ttaatgttaa ctttcaaatc cactagtatt ttttttgctt 2220
ttatgacaaa tagcatacac caaacatttc tgtgaaacta tccttctctt tcaatgtgtt 2280
taattttgga gtaacgtttt ccttgtgact aagttgcaag atcttattta ttaactaggt 2340
atgaagtata aacccatttt ggtgcaatat tcttgactcc ttggtgctaa agattgttaa 2400
attcaatgct tgatgttaca aggtgttgtt aaaacacaaa atgaataaaa gtgagagtag 2460
tcagaactat aacattcaat ttgctattta caaatgaagt atttcatgta atataagtga 2520
acaactggaa ataaagtagg aaagaatttg tatcatgttt tactacatag gttaattttt 2580
taagggatgt tgcaaaggga ttactagaga aagacaaaat gtgaccaaaa aaaagcatga 2640
atatttctta agtatctcaa caacatgtca aagctgcatg tgtaggatgt atgctgtttg 2700
tacaaactat ttcagaatat tttgtaagct ataacatatt tattgtgcat taaaattaaa 2760
tacttttтcc ccaaaggcat gcagtcatga gaattacaga aaatttgcaa catataaagt 2820
agtttgatct aagaggattc aacacctttg ttttgttgct cagtgtgtaa tgactgagat 2880
ttgtaaatct ttgtgaacat tctgtactgg ttcccaagag ctattcattc cctgctacct 2940
gatttcagca caataaatat acttctgctg tgggaaaaat atgatttttt tttgttaatc 3000
agtttattat aatgtaaact gtctgctaga tgtttatttg atttagttaa ttctaaattg 3060
ttttctgctg ctaatttctt atatagaaaa aatcaaacta gaacttgggt atctaccatt 3120
tctggtaaaa agcttatttt ctatagtagc atatttttac tttaaaaatt agctgctttt 3180
ggaagagctt ttcattcaga cataaaacat accattttct tctaaattat tcttaacata 3240
aatggggaag ttcttgtatt ctttcattta aatcattcag tccttactat aagtttcttt 3300
atagaagatg attataagat gtcataaata cagccttttc cagtagcttt cataatttct 3360
ctattatcac agatatgttg catgccttat atctagaaac tgatgaataa agatg 3415

<210> 66
<211> 2289
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1402211CB1

<400> 66
ggacatgccg ggagttgcag taccctcagg aaggtagcgt cttgatctgc gtggcgtggt 60
tctgtgcctt gggaagagat gaatgggaag cggccagcgg agcccggccc agcccgggtg 120
ggaaaaaagg gaaagaagga ggtgatggcg gagttttcgg acgctgttac ggaagaaacc 180
ttgaaaaagc aggtggctga ggcctggagc cgcaggacgc cgttcagtca cgaagtcatt 240
gtcatggaca tggacccttt tcttcactgt gtgatcccaa acttcatcca aagccaagac 300
ttcttagaag ggcttcagaa ggaactgatg aacttggact tccatgagaa gtataatgat 360
ttatataagt tccagcagtc tgatgatttg aagaagagaa gagagcctca catctccact 420
ttaaggaaaa ttctgtttga agatttccgg tcctggcttt ctgatatttc taaaattgac 480
ctggaatcaa ccattgacat gtcctgtgct aaatatgaat tcactgatgc cctgctgtgc 540
catgatgatg agctggaagg gcgccggatt gccttcatcc tgtacctggt tcctccctgg 600
gacaggagca tgggtggtac cctggacctg tacagcattg atgaacactt tcagccgaag 660
cagattgtca agtctcttat cccttcgtgg aacaaactgg ttttctttga agtatctcct 720
gtgtcctttc accaggtgtc tgaagtgctg tctgaagaaa agtcacgttt gtctataagt 780
ggctggtttc atggtccatc attgactcgg cctcccaact actttgaacc ccccatacct 840
cggagccctc acatcccaca agatcatgag attttgtatg attggatcaa ccctacttat 900
ctggacatgg attaccaagt tcaaattcaa gaagagtttg aagaaagttc tgaaattctc 960
ctgaaggagt ttcttaagcc tgagaaattc acgaaagtct gtgaggcctt ggagcatgga 1020
catgtggaat ggagcagccg aggtcccccc aacaaaaggt tttatgagaa agctgaggag 1080
agtaagcttc ctgagatatt gaaggagtgc atgaagttat ttcgctctga ggcactattc 1140
ttgctgctct ccaacttcac aggcctgaag cttcatttct tggccccttc ggaagaagat 1200
gagatgaatg ataaaaaaga ggcagaaacc actgatatca ctgaagaagg gactagccat 1260
agtcctcctg agccagagaa taatcagatg gccatcagca acaacagcca acagagcaat 1320
gagcagacag acccagagcc agaggaaaat gaaacaaaga aagaatcaag tgttcccatg 1380
tgccaagggg aactgaggca ttggaagacc ggtcactaca ctttaattca tgaccatagc 1440
aaggctgaat ttgccctaga cttaattctg tactgtggct gtgaaggctg ggagccagaa 1500
tatggcggtt ttacttctta cattgccaaa ggtgaagatg aagagctgct aacagtgaat 1560
ccagaaagca attctttggc attggtctac agagacagag agactctgaa atttgtcaag 1620
catattaacc accgaagcct ggaacaaaag aaaaccttcc caaacagaac aggtttctgg 1680
gacttttcat tcatctatta tgaatgacag cactgggcaa agctgaacaa aaatgtgacc 1740
cttcgtaatt actgggaagt ctgaaagagc taagcatgga gtcaaggaga actacatggt 1800
agcttgcctg acagtgttct taaaactggt tgtcttttac taggactcat aatgattgtc 1860
ctcaaccgag accttgagct tgcagctaag tacttatctc ttgattaaaa aaaaaaagtt 1920
ggctttttt ttttaacat ttagtccttt ttccatattg gcttcttcag tgaattttta 1980
agttcaattt gtttttattg aggtaaaata tttataacat aaaactgacc agcttaccca 2040
tttttaaata tgcaattcag tggattaagt acattctcat tgttgtccag ccatcaccat 2100
catccatctc cagaagtttt ccatcttccc aaattctgtg cccattgaac aataactccc 2160
cacctcccct tcccctagca acagccatac cttttgtctc tatcatcaac ttcactactc 2220
atatttctca tgtaagtgga atcatacagt atttgtcctt ttgtgactgg tttcacttag 2280
cataaagtc 2289

<210> 67
<211> 4480
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1623474CB1

<400> 67
tacgggagcg tgcgggggcc gcgcgctgct tctctgaggc aggacggcac tgccgggagg 60
cggcggtgac aacgacggcg gtggtgacgg gcaccgggct cgcgggtgag acacagtaac 120
ctggttgaac tctgcatctg gaaagctgaa gactgaagaa agataagaga cattgactag 180

tctggaaaca gggacatctt tggaacttcg ttttcatcca cagtaaactt ttgaagtgtc 240
atcaattgga attgatttct tcatcttatt ctgcctattg ggaagaacat ggcttcaagg 300
attttaagtt tccctttagt tttacatgaa ctttgtagga aacagagccc ttaaagggct 360
tgggaataac aagaagagat tgaagacaga gaagcttgcc ctgttttcct tgccccttca 420
aagaaaagga tttacagctc aaacttagaa cagctgttgt ccagctttag ccatcaagag 480
agaaataaat taaaccacca ttgccagact acaagccctg gtgaagtcag ggtgtgggag 540
tggtggcatt gagaagacta cctaaaagag acaaagactg cagtaaacaa agctcctctt 600
taaagttgga aggggcctca ggttccttct tggattgaaa tagaaataga aacacagggc 660
acacctcttt taggtgcagc tcacatttta tggaactgta gtcgtggagg tactatagta 720
tctcagaaga atttttcttt gcccaaagtt tttttgccat accctgatat tctctccttc 780
ttttgaagac ctgcctccat ccatgagctg tatcttgatc tgtctgactg tccatgtttt 840
ccacctgcaa ccatttgcat gtgtacagcc tactgtttgt ctccagtttt taaactgtac 900
aagttgtgtt tcttaatctt cccttctgcc ttgttctggg gaggtggtta ttcatcattt 960
ggaatcacct ttccccctcc catgtgcttt ccttcatttg agatcttttg acctttggct 1020
ttatttggga gggggaaggg tgataaagtt ttctgtttcc ctggttttct tttgtactcc 1080
tctctgttgc ttccctcctc ccatttttctt gtctgttctg ccgctgtgtg ggcctgggct 1140
atgcggcagg gcagatttcc catcagagct ccaacatgcc cgcagagtct ggaaagagat 1200
tcaaacccag caagtatgtc ccggtctctg cagccgccat cttcctagtg ggagctacga 1260
cactcttctt tgcctttacg tgtccaggac taagcctgta tgtgtcacct gcagtgccca 1320
tctacaatgc aattatgttt ctctttgtgt tggccaactt cagcatggcc accttcatgg 1380
acccagggat tttccctcga gctgaggagg atgaggacaa ggaagatgat ttccgagctc 1440
ccctttacaa aacagtggag ataaagggca tccaggtgcg catgaaatgg tgtgccacct 1500
gccgctttta ccgtcccccct cgatgttccc actgcagtgt ctgtgacaac tgtgtggagg 1560
aatttgatca tcactgcccc tgggtgaata actgtattgg tcgccggaac taccgttatt 1620
ttttcctttt cctcctttcc ctgacagccc acattatggg tgtgtttggc tttggcctcc 1680
tttatgtcct ctaccacata gaggaactct caggggtccg cacggctgtc acaatggcag 1740
taatgtgtgt ggctggctta ttcttcatcc ctgtagctgg cctcacggga tttcacgtgg 1800
ttctggtggc caggggacgc acaaccaatg aacaggttac gggtaaattc cggggaggtg 1860
tgaacccctt caccaatggc tgctgtaaca atgtcagccg tgttctctgc agttctccag 1920
cacccaggta tttggggaga ccaaagaaag agaagacaat tgtaatcaga cctcccttcc 1980
ttcgaccaga agtttcagat gggcagataa ctgtgaagat catggataat ggcatccagg 2040
gagagctgag gagaacaaag tctaagggaa gcctggagat aacagagagc cagtctgcag 2100
atgctgaacc tccacctcct cctaagccag acctgagccg ttacacaggg ttgcgaacac 2160
acctcggcct ggctactaat gaggatagta gcttattggc caaggacagc cccccgacac 2220
ctaccatgta caagtatcgg ccgggttaca gtagcagcag tacgtcagct gccatgccgc 2280
attcctccag cgccaagttg agtcgtgggg acagcttgaa ggagccaacc tcaattgcag 2340
agagcagccg tcaccccagc taccgctcag agcccagctt ggaaccagag agcttccgtt 2400
ctcctacctt tggcaaaagt tttcacttcg atccactatc cagtggctca cgctcctcca 2460
gcctcaagtc agcccagggc acaggctttg agctgggcca gttgcaatcc attcgttcag 2520
agggcaccac ctccacctcc tataagagcc tggccaacca gacacgcaat ggaagcctat 2580
cttatgacag cttgctcaca ccttcagaca gccctgattt tgagtcagtg caggcagggc 2640
ctgagccaga cccaccttta ggctatacct ctcccttcct gtcagccagg ctggcccagc 2700
aacgggaagc tgagaggcac ccacgtttgg tgccaactgg cccaacacac cgagagccct 2760
caccagtccg ttacgacaat ctgtcgcgcc acattgtggc ctctctccag gaacgagaga 2820
agttgctgcg ccagtcaccc ccactcccgg gccgtgagga agaaccaggc ttgggggact 2880
caggcattca gtcaacacca ggctcgggcc atgcccctcg tactagttcc tcctcagatg 2940
attcaaagag atcacctttg ggcaagactc cactgggacg cccagctgtc ccccgttttg 3000
gcaagccaga tgggctaagg ggccggggag tagggtcccc tgaaccaggc ccaacagccc 3060
catacctggg ccgatcgatg tcttacagca gccaaaaagc ccaacctggt gtctctgaga 3120
cagaagaagt ggccttgcag ccattactga cacccaaaga tgaagtacag ctgaagacca 3180
cctacagcaa atccaacggg cagcccaaga gcttaggctc agcctccccct ggcccaggcc 3240
agccacctct cagtagcccc acgaggggag gagtcaagaa ggtgtcaggg gttggtggta 3300
ccacctatga gatttcggtg tgagccttcg gcacctcccc tccccaacgc ctctgcgcct 3360
acaccaaagg gccccaggtg gccaccttcc ttccctcaag gggctccccct cccgtgcatg 3420
gacatttttt aaaccaccga ttccaagagg atgaggagtg ttttctaaaa tgcagtaggc 3480
ttggggagtc ggagagttgg ggccctgaga ctggggtagc aacccccccct tttatctttt 3540
aagaccttcc cttccttgat ccctggacca gactcagtgg acatttgtgc aattgctcgc 3600
cctggaggga accagatcat ttttaaacca gaaataattt tttttattat tgttacggat 3660
tctatttttt tcctcttctg cgttaccagg tgtgtgtgta catataatat atatatatat 3720
atattataaa tatcaaagaa attatatatc tatcctggga tgggaaaatg agggagggat 3780
acatatacgg agggggatct tactcttccc attcctcaga ccagcaggaa aagaggggag 3840
acgtcagtct ttttcctgtg gttccctctc atttgtccca gttactaact acggaaatag 3900
catcctctgc tggtgctaag tgtgattagg aagaagcctg gggagaggtg agtctggaat 3960

tttggtcaca agagggaagg acttggagag gagaattagt tttctaggct cattggcatt 4020
tagtttccct aggaaagggg tcaaaacttc aagacactgg tggtggtggg agatcaggaa 4080
aataacttgg cctagctcaa acaatattgg ataatcccct ccttggggga gagggattag 4140
agtgtgctcc tactggcccc ttggagcctc ccctagctta cacagttaac ttgattttaa 4200
aatccaaggc caggagagaa gaatccaaaa agcaatattt ttcatcacat gccaaaaacg 4260
ggggatagag agaaggagtg gcaggcctag gcccctccga ttgtcccttg ggggttaccc 4320
ctcagcccac ctcactatgg tgctgggtag aggggatacc tgggttctaa cctctaaata 4380
ggggagatcc cagcctccac aaagaggccc ttttattttt tattctgatt agccatttta 4440
aaccaacgag gaataaaaag aaatcctgat ctaaaaaaaa 4480

<210> 68
<211> 1568
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1706443CB1

<400> 68
ctgctcatgt gccccctccc actgaacagg cagccagcgg cctgtgcgag ctcctgtccg 60
tcaacagctg catgggccgt gtgaggcgca tctaccctca gctgctcctg gccctgctca 120
ttcaggtcca ttaccacatc ggcctcaacc tgcctggctg cgtggctcct cccaaggaca 180
ccaagaaggg tgcacagccc tctcccttcg tacctgtgcg ctgggtggtg aaagtggtga 240
aaaccctgct actgaggatg ggctgctctt atgagaccac gtttctggag gaccagggtg 300
gctgggagct catggagcag gtggagagcc accaccgcgg agtggccttg ctggcaaggg 360
ccatggtgca gtactcctgc caggagctgt gccgcatcct ctacctgctc atcccgctcc 420
tggagcgagg cgacgagaag cacaggatca cggccaccgc cttcttcgtg gagctcctcc 480
agatggagca ggtgcgccgg atccccgagg aatactctct ggggcggatg gcagaaggcc 540
tgagccacca cgaccccatc atgaaggtgc tgtccattcg aggcctggtc atcctggccc 600
gcaggtctga gaagaccgcc aaggtgaagg ccctcctgcc ctccatggtg aagggcctga 660
agaacatgga tgggatgctg gtggtggaag cggtccacaa cctcaaggct gtcttcaagg 720
ggcgggacca gaagctgatg gacagtgcgg tctatgtgga gatgctgcag atcctgctgc 780
cgcacttcag cgacgcacga gaggacgtgc gctcctcctg catcaacctg tatgggaagg 840
tggtccagaa gcttcgggca ccacgcactc aggccatgga ggagcagctg gtcagcacct 900
tggtgcccct actgctgacc atgcaggagg gcaactccaa ggtaagccag aagtgtgtga 960
agaccctgtt acgctgttct tacttcatgg cttgggagtt gccaaaaaga gcttatagcc 1020
ggaagccctg ggacaaccaa cagcagacag tggccaaaat ttgcaagtgc cttgtgaaca 1080
cccaccgaga cagcgccttc atattcctca gccagagcct ggagtatgcc aagaactcac 1140
gggcctccct ccggaagtgc tcagtcatgt tcatagggtc cctggtcccc tgcatggaga 1200
gcataatgac agaagatcgt ctgaatgaag tgaaagctgc tctggataac ttgagacatg 1260
acccagaagc atcagtgtgc atctacgcag cccaggtcca ggaccacatc ctggccagct 1320
gctggcagaa ctcctggctg ccgcacggga actcatgggt gtgttactca gccaccaccc 1380
accgctggag ccccagctgt gagaacctgc ccacttccca ccagcggcgc tcctggatca 1440
tgcaggcact gggctcctgg aagatgtcct tgaagaagtg acgtccctga gccccaaacc 1500
ctcctcaggg tggttgagtt ccagccatgc tccctataaa tgtcatgtgg cttaaaaaaa 1560
aaaaaaaa 1568

<210> 69
<211> 1887
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1748627CB1

<400> 69
acgaagtccc acgccccacc cggttctttg ttttttttcaa aataaacata ggttatattt 60
tttcaaacat gagcattcat atattaaaga gcttttcaat ggccgaccat ggtggttcac 120
acctgcaatc ccagcacttt cggaggctga ggcaggtgga tcacttgagg tcaggagttc 180
gagaccagcc tggccagcat ggtgaaaccc tgtctctact aaaaatacaa aaattagtca 240
ggcatggtgg tgcgcacctg tagtcccagc tacttgggag gctgtggcat gagaattgct 300

tgaatccagg aagcggaggt tgcaatgagc ccagatgaca ccactgcact ctagcttggg 360
cgacacagca agactctgtc gaggggcgga gtagcttccg gaaagggtac tgcatttccc 420
gtttctacct ccactgcacc cgcttattgc gtcttgctcc tgggtcacag agcctaaaac 480
gacacaccca acacgcccgc cggagttaca gctaaaggaa ggacagggga agcaatgaaa 540
tgccgagggc ggagccaaga gcgacactgg gggagcagga aaaggcgggg cttccgcttg 600
gggcatggag gctgtacctc ttacgtcact tccgtaaaca aacggagctg cggaggagcg 660
ggtcccggga tgtgaccggg gctctgcttg tggctgcggc ggtggcttct gaggctgtcg 720
ggtctttgcg ggttgcggaa gggggcccca ataccttct tcttcaggtc ttaagaagct 780
ggccgtggtg caataaggaa cttaaaacaa tggaagagcg gaaagtgaag aggaggagtc 840
ctaagtcttt tagtgcccac tgtactcagg ttgtcaatgc caaaaaaaat gccattccag 900
tgagtaaaag cacagggttt tcaaatcctg catcacagtc aacttcacag cgaccaaagt 960
taaaaagagt gatgaaagaa aagaccaaac ctcagggtgg agagggcaaa ggcgctcagt 1020
caactccgat ccagcactcc ttcctcactg atgtctcaga tgttcaggag atggagagag 1080
ggctgctcag tcttttgaat gatttccact ctggaaaact tcaagcattt ggaaatgaat 1140
gttccattga acagatggaa catgttcggg gaatgcagga gaaattagct cgcttgaatt 1200
tggagctcta tggggagtta gaggaacttc ctgaggataa gagaaaaaca gccagtgact 1260
ccaatctgga taggcttctg tcagatttag aagaattgaa ttcttccata caaaaactcc 1320
atttggcaga tgcacaagat gttccaaata cttctgctag ctaaaatgaa atgtagtttg 1380
ctttcttgtg atttgaagag aagcagcagt ctttactttt ccagccaaat ccagtagcag 1440
aatcatcttc cacaacaagg aattaaagta attaaagtgc aagttcaatg atcgttttca 1500
cttactgctt ttagtaaatg tgactcgctg taattcgcca taactggagg cctaggcctt 1560
gttgaaggct tcatcaggaa atgtgacgca gagattgtgc tgctgtgctt catattgttg 1620
ccttatggga ttatacttga aatgcattgt gctgatgttc ttgacagggc ggagggattt 1680
ctctttcctg agctcaccaa acttattcca gtgttgatcg caagctgttg atgcacaggc 1740
gtcttgtggc aagcccagct tcagtttatg ttagacaatc agtacacatg ctggtgtgtt 1800
ttccaatggc cagtgataag aaattaaata agtatttgtg agatttgcta ttttttaata 1860
aagtaattgt tttaaattaa aaaaaaa 1887

<210> 70
<211> 569
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1818332CB1

<400> 70
cggctcgagg gggtggggct gcgggaggcc ctggagcgcg gcggtgatgg cggggccggt 60
gaaggaccgc gaggccttcc agaggctcaa cttcctgtac caggccgccc attgtgtcct 120
tgcccaggac cccgagaacc aggcgctggc gaggttttac tgctacactg agaggaccat 180
tgcgaagcgg ctcgtcttgc ggcgagatcc ctcggtgaag aggactctct gtcgaggctg 240
ctcttccctc ctcgtcccgg gcctcacctg cacccagcgc cagagacgct gcaggggaca 300
gcgctggacc gtacagacct gcctaacatg ccagcgcagc caacgcttcc tcaatgatcc 360
cgggcattta ctctggggag acaggcctga ggcccagctc gggagccaag cagattccaa 420
accactacaa cccttgccaa acacagccca ctccatttca gaccgccttc ctgaggagaa 480
aatgcagact cagggttcca gtaaccagtg atggattcac cccatctccc aaataaagtt 540
tacttgtttt acattccaaa aaaaaaaaa 569

<210> 71
<211> 2338
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1822832CB1

<400> 71
aacgtaggct gtcaggtctc cccctgtgca aatgggacat ccacttgaga gtcaagggtc 60
tgtttgggtg gcagggatag ccacttctga aggtagaaag aaaaataagc caccaaattg 120
gtatctttct gtgaaatgga catcgtgctt agagtctcca ttttccccac aacctggagg 180
aataagtatt gtcatctgca ttttatagct gaggaatctg actcaacaaa attaaattac 240

tcacctaagc aattagcaat taaccaagtc tttctgattc agaaacccag ctgttgcctg 300
ttcatatcca gcccctgtat tgggatcaag atctgcccta ttctcagtgc agcaagatcc 360
aggcagatca cactggactc ccagcactga atctggctca aggggacatg aaatttgact 420
gggtcatggg gctcaggagc atcactctca aaaatagcag tacaggaaga ggcgatggcc 480.
ctaaacagca tttgcaggca gatcccatgt taatcataag ggccaggact ctctcactgt 540
ctgtctctct ctctgtctct cctctagggc tgaccccaca ttggacaccg ctgcatccat 600
gtccatcaca caacacagct gccgtttctt ctgcctgctt atgggaaagt cccctcttct 660
cctccgtttt cttctcttcc tgccctatca caccgtgcac ttctcccttt ccttaaagaa 720
ccaccatcaa ctttaggagg agggaaaggg gtggctctgg caggaaaagc cagaatcccc 780
tctagccagc agagagagag aggaatggct gcatgttttc tcccccagtc caaggcactg 840
ggtcttggct gggttgaagg ttccaagctg ctctcctgct gtgtcggtga gttctggtca 900
acctgcaacc tcctgatatg gccattgcag ttcatcgagt cttcagggac tccccatggc 960
ctggagtact ttgccttgct tacacgggag aggagaatgg atttatagag aacatcatct 1020
aaatccaact tgaccattgt gtggccacac ttgctagatt gctatagtct aaatctagca 1080
ttgtagaaag acgggggagc ttggagctgc acaaacccag gtctggaact ggctccttac 1140
cttggaaggt gaatgatcct gccaggactc ttagcctccc tgggtctcaa tttctttatc 1200
tgtttcatgg gaatgaggat ctctgctggg tggttgggtg atgtgggggc tgtgtgaaaa 1260
cagcttgtca atacaagcca aaatagaaat atttctccac agagtatgaa ggtcaaatga 1320
gagaatacat ttaaattaaa tggaaaatta aaatggtaaa aaatgcaaag ctgtattgaa 1380
agttccgagc ttctctataa ggagcttttt gactatgtaa gaatcctgta ctcgttcccc 1440
ctaaatataa aaaaaagttg aaggaggcag aagggagagt gatgcacgat gggcaaggac 1500
ttcacctgct gttgctggct ttgaggatgg aggaaggagg ccacaaacct agaagctgga 1560
gcccctagaa gctagaaaag gcagggaccc gattctttcc ttgagcctcc agaagggaca 1620
cagccccgcc agcaccttga ctttagcctg gtgagatcct cttaggactt ttggcaatca 1680
gaactataag acagaaatgg aagccactga gtctgtagct gtttgttgca gcagcaatag 1740
aaaactaatg cagggcccaa gaaatcactg gtgatgagat cgggaaagtg ggctcaggag 1800
gtctggatct gtgatgagat ggggaaagtg ggctcaagag gtctggatct gtggtgagat 1860
gggggaagtg ggctcaggag gtctggatct gtgatgagat ggggaaagtg ggctcaggag 1920
gtctggatct gtgatgagat ggggaaagtg ggctcaggag gtctggatct gtgatgagat 1980
gggggaagtg ggctcaggag gtctggatct gtgatgagat gggggaagtg ggctcaggag 2040
gtctggatct gtgatgagat gggggaagtg ggctcaggag gtctggatct gtgatgagat 2100
ggggaaagtg ggctcaggag gtctggatct gtgatgagat gggggaagtg ggctcaggag 2160
gtctggatct ggggtgggga tctggagtgg aaggggaatt catttgttca ttgtctatcc 2220
ttttgtattg attgaatttt ttatatatat atgtgaattt tcacaataaa attttttcc 2280
aaaataaaat aaacaaaagg ggctttttgc aacccaattc ctatctaaaa aaaaaaaa 2338

<210> 72
<211> 481
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1832219CB1

<400> 72
cctgggcggc gttggtccgg tgcgtcctgt tctacagcta tggccgggcc agctgcagct 60
ttccgccgct tgggcgcctt gtccggagct gcggccttag gcttcgcttc ctacggggcg 120
cacggcgccc aattcccaga tgcctacggg aaggagctgt ttgacaaggc caacaaacac 180
cacttcttac acagcctggc cctgttaggg gtgccccatt gcagaaagcc actctgggct 240
gggttattgc tagcttccgg aacgaccttа ttctgcacca gcttttacta ccaggctctg 300
agtggagacc ccagcatcca gactttggcc cctgcgggag ggaccctgct actcttgggc 360
tggcttgcct tggctctttg agctcccttt tgcttaatta ctgggttttc tgggcagttt 420
ttttttaaa gagttggagt aagaagagga ttaaaaagga aaggcaaata aaaaaaaaaa 480
a 481

<210> 73
<211> 1255
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature

<223> Incyte ID No: 1899010CB1

<400> 73
cggacggtgg gcggacgcgt gtgctgcggc gtcctagctg gcttacaggg cggcggcggg 60
gtgtgtgtcc tctgttaaga gtgctactcg cccggggttg atctgtgcat gccactcctg 120
ggtcagacgg tgaggtcggc gtctgcgagg acgcggcggt ggagtagaag ggcagccgga 180
gacaggcccg gcgccccttc cgaggctaga cggccccagc ttcgcgggga tcatggcatt 240
gctggtggac cgagtgcggg gccactggcg aatcgccgcc gggctcctgt tcaacctgct 300
ggtgtccatc tgcattgtgt tcctcaacaa atggatttat gtgtaccacg gcttccccaa 360
catgagcctg accctggtgc acttcgtggt cacctggctg ggcttgtata tctgccagaa 420
gctggacatc tttgccccca aaagtctgcc gccctccagg ctcctcctcc tggccctcag 480
cttctgtggc tttgtggtct tcactaacct ttctctgcag aacaacacca taggcaccta 540
tcagctggcc aaggccatga ccacgccggt gatcatagcc atccagacct tctgctacca 600
gaaaaccttc tccaccagaa tccagctcac gctgattcct ataactttag gtgtaatcct 660
aaattcttat tacgatgtga agtttaattt ccttggaatg gtgtttgctg ctcttggtgt 720
tttagttaca tccctttatc aagtgtgggt aggagccaaa cagcatgaat tacaagtgaa 780
ctcaatgcag ctgctgtact accaggctcc gatgtcatct gccatgttgc tggttgctgt 840
gcccttcttt gagccagtgt ttggagaagg aggaatattt ggtccctggt cagtttctgc 900
tttgcttatg gtgctgctat ctggagtaat agctttcatg gtgaacttat caatttattg 960
gatcattggg aacacttcac ctgtcaccta taacatgttc ggacacttca agttctgcat 1020
tactttattc ggaggatatg ttttatttaa ggatccactg tccattaatc aggcccttgg 1080
cattttatgt acattatttg gcattctcgc ctatacccac tttaagctca gtgaacagga 1140
aggaagtagg agtaaactgg cacaacgtcc ttaattgggt ttttgtggag aaaagaatgt 1200
tgtcccaaga agataaaaaa tattgttaag tgtgcaagtt attaaaaaaa aaaaa 1255

<210> 74
<211> 875
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2008768CB1

<400> 74
tagctgagca cgccctctga gccgctcggt ggacaccagg cactctagta ggcctggcct 60
acccagaaac agcaggagag agaagaaaca ggccagctgt gagaagccaa ggacaccgag 120
tcagtcatgg cacctaaggc ggcaaagggg gccaagccag agccagcacc agctccacct 180
ccacccgggg ccaaacccga ggaagacaag aaggacggta aggagccatc ggacaaacct 240
caaaaggcgg tgcaggacca taaggagcca tcggacaaac ctcaaaaggc ggtgcagccc 300
aagcacgaag tgggcacgag gaggggtgt cgccgctacc ggtgggaatt aaaagacagc 360
aataaagagt tctggctctt ggggcacgct gagatcaaga ttcggagttt ggacctcttc 420
aacgacctga ttgcttgtgc gttccttgtg ggagccgtgg tctttgctgt gagaagtcgg 480
cgatccatga atctccacta cttacttgct gtgatcctta ttggtgcggc tggagttttt 540
gcttttatcg atgtgtgtct tcaaagaaac cacttcagag gcaagaaggc caaaaagcat 600
atgctggttc ctcctccagg aaaggaaaaa ggaccccagc agggcaaggg accagaaccc 660
gccaagccac cagaacctgg caagccacca gggccagcaa agggaaagaa atgacttgga 720
ggaggctcct ggtgtctgaa acggcagtgt attttacagc aatatgtttc cactctcttc 780
cttgtcttct ttctggaatg gttttctttt ccattttcat taccaccttt gcttggaaaa 840
gaatggatta atggattcta aaagcctaaa aaaaa 875

<210> 75
<211> 2188
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2070984CB1

<400> 75
cggcgacggc gacggcagcg gggacggcag cagtagcggg agcagcagcg tggacgcggc 60
tggcgctggc gccatgaacc cgctgtaagg cgcaggctgt gcagcacggg gtgcgggga 120

ggaggaggag gacgccgcgg tgaagttctc cgccatgaac ctgaggggcc tcttccagga 180
cttcaacccg agtaaattcc tcatctatgc ctgtctgctg ctgttctctg tgctgctggc 240
ccttcgtttg gatggcatca tacagtggag ttactgggct gtctttgctc caatatggct 300
gtggaagtta atggtcattg ttggagcctc agttggaact ggagtctggg cacgaaatcc 360
tcaatatcga gcagaaggag aaacgtgtgt ggagtttaaa gccatgttga ttgcagtggg 420
catccacttg ctcttgttga tgtttgaagt tctggtctgt gacagaatcg agagaggaag 480
ccatttctgg ctcctggtct tcatgccgct gttctttgtt tccccggtgt ctgttgcagc 540
ttgcgtttgg ggctttcgac atgacaggtc actagagtta gaaatcctgt gttctgtcaa 600
cattctccag tttatattca ttgccttaag actggacaag atcatccact ggccctggct 660
tgttgtgtgt gtcccgctgt ggattctcat gtcctttctg tgcctggtgg tcctctacta 720
cattgtgtgg tccgtcttgt tcttgcgctc tatggatgtg attgcggagc agcgcaggac 780
acacataacc atggccctga gctggatgac catcgtcgtg ccccttctta catttgagat 840
tctgctggtt cacaaactgg atggccacaa cgccttctcc tgcatcccga tctttgtccc 900
cctttggctc tcgttgatca cgctgatggc aaccacattt ggacagaagg gaggaaacca 960
ctggtggttt ggtatccgca aagatttctg tcagtttctg cttgaaatct tcccatttct 1020
acgagaatat ggaaacattt cctatgatct ccatcacgaa gataatgaag aaaccgaaga 1080
gaccccagtt ccggagcccc ctaaaatcgc acccatgttt cgaaagaagg ccagggtggt 1140
cattacccag agccctggga agtatgtgct cccacctccc aaattaaata tcgaaatgcc 1200
agattagatg ccacttccgg ggacagagct taagtggact gggacgcact ctctccgcct 1260
tcctctgccc cctcgttcac cccgcagacc agaaccagta ctggagctgg gtctccaggt 1320
acgtccatct catgccttgt ttgcatccag cgcctatcag ccactcacca cgacgggacg 1380
cggaagtggc aggtgacggg ggtgtgtgcc agcagatgcg gatgccagga agagtgtgag 1440
aacaggggtg ggattaccgt ctgtctggga ggggctccag gtacccctct tccccgtcag 1500
acccactggg agatggctgc ttgccaggcc cccagaagga acatctgtct atacggtgct 1560
gaaatcccaa tcaaaagtat tgtttagaaa tgtatttctc cacagggctg acctcctgca 1620
gctcgctgag cactcccagg tcctcagcac tcccaggtcg tggctggggc agtcagtagg 1680
aactgtaact atgtctctga tgcaccacgt gtttagacac agcacagtcc ttttttctgt 1740
tcctactgtg gaagtagttt ctctttgggc atgctgacag cagtttttca tagcctcacg 1800
gatgagccct ttctacggga gtgactccat gcttgtatac agagtattta tacaaatgtt 1860
ttagcatctt catatgcggt gttaacccct agttctgtac agcatattct gttcaagtat 1920
ttttttacaa gcttgtgctg taggcacatg ccttctgctg cagaagtgga cgcccgtggc 1980
acactccccc ccccccccg tggggtgcca cgccttcatg ggacattgcc acttctgccc 2040
tggaactcgt gcaggtacgt agtagctgct actgccacaa cggcaacacc aagcaagaga 2100
tggtccatgc ttttctgacg ttctcagaat agtggctagc ttcaaacctg acaagcgctg 2160
cttgaagccg gaacactaga gaatgttg 2188

<210> 76
<211> 1561
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2193240CB1

<400> 76
tctggccatg ctcctgagga aagctggtgc catcacactc ccgtctgtta ccgtggccct 60
ggccaagcac tggacagcgg cgattgtgct atgaagggaa atcctccaaa gatatcctga 120
aaagagtagc tgcaaatgca ttgatgtcac tgctggctgt cagtagaaga gcacagaaac 180
atgctttgaa agccaatctt atagacaatt gcatggagca gatgaaacac ataaatgcac 240
aactgaacct agattctctg aggcctggga aagcagcatt gaaaaaaaag gaggatggtg 300
ttattaaaga gttaagcatt gccatgcagc tcctaagaaa ctgtctttat caaaatgagg 360
aatgtaaaga agcagctctt gaagctcacc ttgtccctgt cttgcactct ctctggcctt 420
ggattttgat ggatgattca ttgatgcaaa tttctctgca gctcctttgt gtctatactg 480
caaattttcc aaatggttgc agttctcttt gttggtcaag ttgtggacaa caccctgttc 540
aagctacaca tagaggagcc gtgagcaact ctctgatgct gtgtatccta aagttggctt 600
cccagatgcc actggagaac accacggttc agcagatggt ttttatgctt ctttcaaacc 660
tggccttgtc gcatgactgt aaaggagtaa ttcagaagag taacttctta cagaacttcc 720
tctctctagc attgccaaaa ggaggaaata aacatctaag taatctgact attctttggt 780
tgaagttact cctgaatata tcatctggag aagatgggca acaaatgatt ctgaggcttg 840
atggctgtct agacttacta acagagatga gcaaatacaa gcacaagagc agccctttat 900
tgcctcttct tatctttcat aatgtttgct tcagtcctgc aaataaaccc aagatcctgg 960
ctaatgaaaa agtcattact gtgcttgctg cctgtctgga aagtgagaat caaaatgctc 1020

agaggattgg agcagctgcc ctttgggctc tgatttacaa ttatcagaag gcaaaaacag 1080
ctttgaaaag cccatcagta aaaagaagag tggatgaagc atactcctta gcaaagaaaa 1140
ctttcccaaa ctcagaagca aaccctctaa atgcctatta tttgaaatgt cttgaaaacc 1200
tcgtgcagct ccttaattct tcctgagtgc catgggatgc tacaccttga agctgacagt 1260
catcaacagg ggagctaaag ttgaagccag ctgtgtgtag cagctgttac ctgaagacgt 1320
gctacctctc tacaaagtgt tgatcccctt ctttcccatg agagagagaa ctggtgatac 1380
tccaacaccg tccagttgtg gcagctctcc agaagtaata gcagctgaca actttctgtg 1440
ccttttcctt tctgttgaaa aggcatagaa agttctggga acataaacat ttttaccctt 1500
ttctatgcca tttattttgt aaaaatccta tttaacagtt atttaataaa acaagatttt 1560
g 1561

<210> 77
<211> 1777
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2235177CB1

<400> 77
ccaacatcag cgagggaggc agcttcgggg agctggggac catgggctcc aggatcaagc 60
agaatccaga gaccacattt gaagtatatg ttgaagtggc ctatcccagg acaggtggca 120
ctctttcaga tcctgaggtg cagaggcaat tcccggagga ctacagtgac caggaagttc 180
tacagacttt gaccaagttt tgtttccccc tctatgtgga cagcctcaca gttagccaag 240
ttggccagaa cttcacgttc gtgctcactg acattgacag caaacagaga ttcgggttct 300
gccgcttatc ttcaggagcg aagagctgct tctgtatctt aagctatctc ccctggttcg 360
aggtatttta taagctgctt aacatcctgg cagattacac gacaaaaaga caggaaaatc 420
agtggaatga gcttcttgaa actctgcaca aacttcccat ccctgaccca ggagtgtctg 480
tccatctcag cgtgcattct tattttactg tgcctgatac cagagaactt cccagcatac 540
ctgagaatag aaatctgaca gaatattttg tggctgtgga tgttaacaac atgttgcatc 600
tgtacgccag tatgctgtac gaacgccgga tactcatcat ttgcagcaaa ctcagcactc 660
tgactgcctg catccacggg tctgcggcga tgctctaccc catgtactgg cagcacgtgt 720
acatccccgt gctgccgccg catctgctgg actactgctg tgctcccatg ccctacctca 780
taggaatcca tttaagttta atggagaaag tcagaaacat ggccctggat gatgtcgtga 840
tcctgaatgt ggacaccaac accctggaaa ccccccttcga tgacctccag agcctcccaa 900
acgacgtgat ctcttccctg aagaacaggc tgaaaaaggt ctccacaacc actggggatg 960
gtgtggccag agcgttcctc aaggcccagg ctgctttctt cggtagctac cgaaacgctc 1020
tgaaaatcga gccggaggag ccgatcactt tctgtgagga agccttcgtg tcccactacc 1080
gctccggagc catgaggcag ttcctgcaga acgccacaca gctgcagctc ttcaagcagt 1140
ttattgatgg tcgattagat cttctcaatt ccggcgaagg tttcagtgat gtttttgaag 1200
aggaaatcaa catgggcgag tacgctggca gtgacaaact gtaccatcag tggctctcca 1260
ctgtccggaa aggaagtgga gcaattctga atactgtaaa gaccaaagca aatccggcca 1320
tgaagactgt ctacaagttc gcaaaagatc atgcaaaaat gggaataaaa gaggtgaaaa 1380
accgcttgaa gcaaaaggac attgccgaga atggctgcgc ccccacccca gaagagcagc 1440
tgccaaagac tgcaccgtcc ccactggtgg aggccaagga ccccaagctc cgagaagacc 1500
ggcggccaat cacagtccac tttggacagg tgcgcccacc tcgtccacat gttgttaaga 1560
gaccaaagag caacatcgca gtggaaggcc ggaggacgtc tgtgccgagc cctgagcaaa 1620
acaccattgc aacaccagct acactccaca tcctacagaa aagcattacc cattttgcgg 1680
ccaagttccc gacgagaggc tggacctctt catcacattg acttacgccg ttgcttttcc 1740
agactgggca gaggggctga cttcgcagtg tgtgcca 1777

<210> 78
<211> 1841
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2416227CB1

<400> 78
cgaaagagaa acccggaggg cgccggggac tgggccgggg tctgcagggc tcagctgagc 60

ccatgagctc ccagagctaa cccctgaaca cccaggcggg caaagggctg atgtcggtag 120
tccccatcct ggaggggcag gctctgcgca tctgctcctg gcatggcgct gcggcacctc 180
gccctcctgg ctggccttct cgtgggagtc gccagcaagt ccatggagaa cacggcccag 240
ctgcccgagt gctgtgtgga tgtggtgggc gtcaacgcca gctgcccagg cgcaagtctg 300
tgtggtccag gctgttacag gcgctggaac gcggacggga gcgccagctg cgtccgctgt 360
gggaacggaa ccctcccagc ctacaacggc tccgagtgta gaagctttgc tggcccgggt 420
gcgccattcc ccatgaacag aagctcaggg acccccgggc ggccacatcc tggggctccg 480
cgcgtggccg cctccctctt cctgggcacg ttcttcatta gctccggcct catcctctcc 540
gtagctgggt tcttctacct caagcgctcc agtaaactcc ccagggcctg ctacagaaga 600
aacaaagctc cggccctgca gcctggcgaa gccgctgcaa tgatcccccc gccacagtcc 660
tcagtacgga agccgcgcta cgtcaggcgg gagcggcccc tggacagggc cacggatccc 720
gctgccttcc cgggggaggc ccgtatcagc aatgtctgac ctggaggccg agaccacgcc 780
acgcacttgg cggcagggac ccggaggccg acccccttggc gggaaccagc acaaagtgtt 840
ggcatcgccc ggcgcccggg acagtcctgg gcacagcctc ggctctgagt ccctccgcct 900
cccagcgacg gacgccaaag ggtcccgggc cgcctgaggc tcctccccac cacagccatc 960
tcgtttatcg gaccaggagc aggcatccat gagacctcag agcttcagat cgaggccttg 1020
gggggtccgg gccccccccag gaaacacggt gaggccccag cgcctgcagc caaagctggc 1080
acgatctatg gggcaggtgc cgctctgcct agaaaagcca ggggctctgc tgccgtgccc 1140
tccagagccc acagcgggca ggactcctcc agcaccacca cacccagtgg cccgagaccc 1200
ctctgagaac agtgaggctg gtcctcgtgc cgttccagcc ggtgcccggc cagtggggag 1260
gacacagcct aggaaccagc tgcctgagac cagggtgcct ctgggctgtc ctcccgcgtg 1320
gcggagaccc caagcacgca gccacccatt tccggagctg caggatagag cttcctcttg 1380
atctctgttt ttaagcagaa attcattgtg cagaaaagtc ctccagagct ctgtggcccc 1440
gctcggatcc gctggacccc catgcctggc tgatccctgc ccacgtgggg caggcccaca 1500
tctaaccccc acaagtcact gcctcactgc acctgccaag gctgccctgg cgctgagtcc 1560
tggggtccct cccggagttc ctgggagaaa ggcgccgtcg tggccgcctc ccgcacgcca 1620
ggcccgggct ccaccgtggg tctcagacgc cctgcggcac cggcaccgtc tgctttagca 1680
tgggaccccc atctgagggg tggcctggcc ttcggggtcc ccacgctcct ttgcgaagtc 1740
cactgtgggt gccatcatgg tctccgggac ctgggccagc gggaacgtgg gggcactggg 1800
tgtgctgata taaagtcggc attactcaaa aaaaaaaaaa a 1841

<210> 79
<211> 1616
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2461076CB1

<400> 79
agctggagac agagaggctg gccattgcgg gcataagcag aggcctagat caacagtgcc 60
aaggactggg atgagggtag tggtggtggg agtggagaac catgatgatg taagagacag 120
ttagtggctt ctgatccagg gaaggtggca gtgtagtgac agaaacctga aggggagtgg 180
atgggagtga tggccccact gcctgggggg agaggccagg ctggaggcgg gacctggagg 240
aatggaggcc agtgctaccc tcacactccc agcctcctgt cctggtggga tttggaccca 300
cagatcccca tggtgcatat gtcaagtctg gcctgttcca cctctttctc aaaggaaggt 360
tctgagcctc cctacccagt gtaggggtgg caacgatccc tcattcattc aagcatttat 420
gaaccaccct ccctaagtct agctcatgca gcaccaaatt tattcactta gctaatgatt 480
gacctactgc tgtgttctag gcaccaggga taccacatgg aggcgtgcgc cagagccctg 540
tcccctggag cccttgggtt caaatactgg ttctacctgg gagtgggcag tagttggcct 600
cccagccttg catcctcagc cattgaatta gttaacatgt gaagactctt agaacagcgt 660
aagcgctgca gaagaatttg ctgctgctgt tattattatt ccagacatta ttataaacag 720
aaagccgttc tgagctcaga agcgaactct cagcgggtcc ttctctaccc agaagaggtg 780
ttcataggaa cctgttagtc agagtgggaa cgggccaagg gtcactaggt cactgttttt 840
ttgttggttt ttttaagttt tataattaat ataaaaatgg ggtctcgcga tgttggccat 900
gttgttcttg aacttttggc ctcaagcagt cctcccgcct cagcctccca aagtgctagg 960
attacgggca cgagccacca catccagcca agctgggtca ctttgaactg aggccagggt 1020
ggattgcaga cgtggggcac atcccagccc ctgaattcag ggtccaccat cacacctgct 1080
cctgcctagc atcctgagag ggctgcccag aaaagagctt ttacaaacca gctgttgata 1140
agctcagaat tttagaaacc actgtcttca ttataagggc ataaggtaca aggatgaaac 1200
tccccctgtc tttactcttt cttagaactc tgggttttta catcccagtg aaaggggatc 1260
tgtcctctgg ttgtgaggac aaggcatgtc tttatgtgtt aaaaagagta accactgaca 1320

aggtttttt tgatcctttt aaaatttact ttcgacctgt aatcccagga ctttgggagg 1380
ccgaggcagg tggatcactt gggttaggag tttgagacca gcctggtcaa catggtgaaa 1440
ccccgtctct actaaaaaat acaaaaatta gctgggtgtg gtggtgtgca cctggaatct 1500
cagctattcg ggaggctaag acaggagaat cgcttgaacc caggaggcag aggttgcagt 1560
gagccgagat cacgccactg cactctagcc tgggtgacag agcgagactc catgtc 1616

<210> 80
<211> 1434
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1957517CB1

<400> 80
ccgccccgtc cgttgagggc ccgcgccgca tggaggccgg ctgaggagcg ccgctgcctc 60
gcctcggctc cccacaggtg caggaagccg ccgcccagcc atggacgggg aggagcagca 120
gccaccgcac gaggccaacg tggaacctgt tgtgccgtca gaggcttcag agccggtgcc 180
cagggtgctt tctggagacc cccagaacct gtccgacgtg gacgccttca acctgctcct 240
ggagatgaag ctgaagcggc ggcgtcagcg gcccaacctg ccgcgcactg tgacccagtt 300
ggtggctgag gacgggagca gggtgtacgt ggtggggaca gcccacttca gcgacgacag 360
caagagggac gttgtgaaga ccatccggga ggtgcagcct gacgtggtgg tcgtggagct 420
ctgccaatat cgtgtgtcca tgctgaagat ggacgagagc acgctgctgc gggaggccca 480
ggagctcagc ctggagaagc tgcagcaggc cgtgaggcag aacgggctca tgtcggggct 540
gatgcagatg ctgctgctga aggtgtctgc acacatcacc gagcagctgg gcatggcccc 600
aggtggcgag ttcagggagg ccttcaagga ggccagcaag gtgcctttct gcaagttcca 660
cctgggtgac cgacccatcc ccgtcacctt caagagggcc atcgcagcgc tctccttctg 720
gcagaaggtc aggctggctt ggggcctgtg cttcctgtca gaccccatca gcaaggatga 780
cgtggaacgc tgcaagcaga aggacctact ggagcagatg atggccgaga tgattggcga 840
gttcccagac ctgcaccgca ccatcgtctc ggagcgcgac gtctacctaa cctacatgct 900
gcgccaggcc gcgcggcgcc tcgagctgcc tcgggcctct gacgccgagc ccaggaagtg 960
cgtcccctcc gtggtcgtgg gcgtcgtggg catgggccac gtgcctggca tcgagaagaa 1020
ctggagcacc gacctcaaca tccaggagat catgaccgtg cccccgccgt ccgtctccgg 1080
cagagtgtct cggttggccg tgaaggccgc cttcttcggc ctgctgggct acagcctgta 1140
ctggatgggc cgccgcaccg cgagcctggt cctgtcgctg cccgccgcgc agtactgcct 1200
gcagagggtg accgaggccc ggcacaagta ggagactgct ccccgcccgc tcgggcccct 1260
gaggagccag tgcccccgcg gcacttctgg gtgccaggtg catcctagcc cgcccgaggc 1320
ccctgccacc ccccatgggg gtctgggccc ggcctcgcct gccctcctgg gccagtcacc 1380
cctcccccag cccacccaaa taaaggatta tttaactgtc tgaaaaaaaa aaaa 1434

<210> 81
<211> 2085
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 866038CB1

<400> 81
ccgcgcccgg ccccgccatg gtgtcctgga tcatctctcg cctggtggtg ggctgatgcg 60
gagctgggag ggagaggcca ctgcccttgg ctgagaggcc cagtaaccat gcctgcctcc 120
ctaggctcat ctttggcacc ctgtacccag cctattcttc ctacaaggcc gtgaagacaa 180
aaaacgtgaa ggaatatgtg aaatggatga tgtactggat cgtctttgcc ttcttcacca 240
cggccgagac gctcacggat atagtgctct cctggttccc cttctacttt gaactgaaga 300
tcgccttcgt gatatggctg ctgtcccctt acaccaaggg ctccagcgtg ctctaccgca 360
agttcgtgca cccaacgctg tccaacaagg agaaggagat cgacgagtac atcacgcagg 420
cccgagacaa gagctatgag accatgatga gggtgggcaa gaggggcctg aaccttgccg 480
ccaatgctgc agtcacagct gccgccaagg gccagggggt gctgtcagag aagctccgca 540
gcttcagcat gcaggacctg accctgatcc gggacgagga cgcactgccc ctgcagaggc 600
ctgacggccg cctccgaccc agccctggca gcctcctgga caccatcgag gacttaggag 660
atgaccctgc cctgagtcta aggtccagca caaacccggc agattcccgg acagaggctt 720

ctgaggatga catgggagac aaagctccca agagggccaa acccatcaaa aaagcgccca 780
aagctgagcc actggcttcc aagacactga agacccggcc caagaagaag acctctggcg 840
ggggcgactc agcttgagcc cctccacccc cgcaggctgc agagcaagga tgaagcctca 900
ggaggggcct cagacccagc ccctgctcca cactgtgcca gtagcctagg tgtctcaggc 960
ccctgggccc cgcagatggc catttccggt gcctgcccag tggccactct tctggaaggg 1020
gcttggaaaa gaggaaggag gcccagctgt gggggttgag ggtagagggt ggaccagagg 1080
ctgaggactg agccacccaa ggaggtgggg actgctcggc ctccaccgct gttccgctgc 1140
agccccgccc tgccccaccc acccagtgcc ttgctgaagc ccatagcaat ccgcttctca 1200
gaggtcctat cgtgtttcca ctgctcgcct tggtttggga gcagggaggg ggaagtccta 1260
gcccagatgg accaaggacg ggcctgaagg cacatggggg aaagggagca cacggggagg 1320
acgttgggga ccctgggtgg ggcctccagg tgcagctgtg gatggaagac agggattggc 1380
ctgtgcttca gcgaccagga tggccaggcc agagctgcag ctgggggctc ttttcctggt 1440
cattgggtgg ggctgagtgc cacatgttcc cacattaaaa aggggggtcc agggctgtgt 1500
gagtgtgtct ttctgggtct agggctcggg gtagtttggg tcaaggactg tccctccagc 1560
agtcgcctcc tcccaccctg agccccacag tcatctggcc ctttccctgc tcaaccctcc 1620
atcctaggct ctgagcctca gaggacccag cccatgagag aacggggatc tgggggggcct 1680
ctcacctgct cctatgacct tgctcccttt taggtcaccc cattgccacc gtgcccctgg 1740
gctggactcc cgtgctcctc agggcccacc cctgctctgt ctggtacagg cccctgctga 1800
gtgggccccc ctcctctgcc cctggggtcc atcccctctt gcccagggtc cccatcctgt 1860
accaagcaga ctgggcccta agacccctgg cagaacccag cctctgctca cacccgcccc 1920
agcttctgcc acggcttcag tcagggccag gaggaacacg tgaaggagaa agagaaatgc 1980
aggagccgcg gggctcccgg ttccttggga agagggtgcc cattggacct ttggcactgg 2040
atgagccaat aaaccaaact ctggcacctc aaaaaaaaaa aaaaa 2085

<210> 82
<211> 904
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3869704CB1

<400> 82
ggggagccca gctgtgctgt gggctcagga ggcagagctc tgggaatctc accatggcct 60
ggacccctct cctgctcccc ctcctcactt tctgcacagt ctctgaggcc tcctatgagc 120
tgacacagcc accctcggtg tcagtgtccc caggacaaac ggccaggatc acctgctctg 180
gagatgcatt gccaaaaaaa tatgcttatt ggtaccagca gaagtcaggc caggcccctg 240
tgctggtcat ctatgaggac aacaaacgac cctccgggat ccctgagaga ttctttggct 300
ccagctcagg gacaatggcc accttgacta tcagtggggc ccaggtggag gatgaagctg 360
actactactg ttactcaaca gacagcagtg gtaatgatag ggtgttcggc ggagggacca 420
agctgaccgt cctaggtcag cccaaggctg ccccctcggt cactctgttc ccaccctcct 480
ctgaggagct tcaagccaac aaggccacac tggtgtgtct cataagtgac ttctacccgg 540
gagccgtgac agtggcctgg aaggcagata gcagccccgt caaggcggga gtggagacca 600
ccacaccctc caaacaaagc aacaacaagt acgcggccag cagctacctg agcctgacgc 660
ctgagcagtg gaagtcccac aaaagctaca gctgccaggt cacgcatgaa gggagcaccg 720
tggagaagac agtggcccct acagaatgtt cataggttct catccctcac cccccaccac 780
gggagactag agctgcagga tcccagggga ggggtctctc ctcccacccc aaggcatcaa 840
gcccttctcc ctgcactcaa taaaccctca ataaatattc tcattgtcaa tcagaaaaaa 900
aaaa 904

<210> 83
<211> 1496
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1415179CB1

<400> 83
gacgacagaa gtttgtacgg ttccaggagc tggctgtgca ggctcctcgg gccttctgct 60
acatgcctgc tggggacgcc cagctactcc tggcccccag cttcaaggga cagacgctgg 120

tgtatagaca cattgtggtg gatctcagtg cctaggggt gccgaggcct ctggtctcct 180
cagggtggcc actggaggat gggtgggggc ctctgtgagc accacgtatg cctgtgtgaa 240
gacaagcgac caacctgcat tcatgccctt gtatccgcac acgcccagtg tgtacaggcg 300
caccccgtgg actggcccag gaggacgtgc atcatcactg caatcagcat gaacaccagg 360
caccctgggg acacctaacg cccagttccc gtgaccccaa ctgcaggccc cccctactct 420
gcagcctttc ccatgctctg ccccctgtct catgcacgag tggtgcgggg tgctgggagc 480
gaggggggcc gcatcctcct gtccttgtgc ttctccttct gcccctctgg tctttcctgt 540
tggtgctcca ggcattgttt acctgctctt gctccccgct gtagcccaca gccatatctg 600
tcatgcttcc ccggggccac tcacccctgc cccaccccgt ctgcatgctc acatggacgc 660
ggacggacac actcattaca cactcacacg ccccgactgc acccggtgtc catctataaa 720
catgtcaggg cacgtgtgca tacaagcaga ttcagcactg cctaccaggc tctcctgtta 780
ccttgcctct ctgcttggag ggggcccccc ctgctcaccc ccagcgtgcc ccccccggaa 840
ctgataagga tgagaatggt ggtgccagct tctgagggcc tgcttggcct cctgggggca 900
aagcccctct gcccgaagca ataacaacag cagcagaagc aatcccgggg agctgccagg 960
ggtgtctgtc tctgttctcc ttggcgtttc cttcatggcc gtgaatccca gctcacccat 1020
ggacagaaga cctcccctca tctgaaagcc tttctgcctt ccttcctgta gctgtcgctt 1080
tctcaggctt ggactgtccc ctggttcacc ctgtctttgg gatgtggccc ttgagctggg 1140
cgggctcctt gggtggatta gagagctggc cgagctgcaa gcacctgagt tacctgcgag 1200
ggtgcccaga gtggagcttg gcctgagggc gtgttgcagc cctggcccca ccagtccaca 1260
gaggccccca ggagcactgc ctttcagtgg accgggtggg ggctggtggg aggggcccca 1320
ctagctgagc ttctgctgcc ctctgctgtc tgctgagatg ccaggctagt ggagagggca 1380
caccatgtgc ccaaatggtg tgatgctcct tgtagcagga taccagggta ctgggggca 1440
atgctatgat taacttgctt caaataaaaa gttcccgccc caaaaaaaaa aaaaaa 1496

<210> 84
<211> 2837
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1664792CB1

<400> 84
gctgcagctg cgccgccgcc ctggcgcagg ccacgcccga ccacttccac atgcagtcgg 60
agtctctgat cgggcccctg atgcagacca tctcccacca gcactggaag gtccgtgtgg 120
ccgccattga agccacaggc gcagtgatcc attttggcaa cgggaagtcc gtggacgacg 180
tgctttccca ttttgctcag cgactgtttg atgacgtccc gcaggtccgg cgggcggtgg 240
cctccgtggt gggcggctgg ctgctgtgtc tgcgtgaccg ttactccttc ttccacaagc 300
tcatccctct gctgctcagt agcctcaacg acgaggtgcc tgaggtcagg cagctggctg 360
ccagcctctg ggaggacgtt ggcctgcagt ggcagaagga gaatgaggag gacctgaagg 420
acaagctgga ctttgccccc cccaccccac cccattaccc tccacatgag cgccgccctg 480
tgctgggctg ccgggagctc gtcttcagga acctctccaa gatcctccct gccctgtgcc 540
acgacatcac cgactgggtg gtggggaccc gagtgaagtc ggcacagctg ctcccagtgc 600
tgctgctgca tgccgaggac cacgccacgc agcacctgga ggtcgtcctc cggaccctgt 660
tccaggcctg caccgacgag gaggcagccg tggtccaaag ttgtaccaga tctgcagagc 720
tcgtcgggac gtttgtcagc cctgaggtgt ttctgaagct gatcttatcg acgctgaaga 780
agacgccctc tgcctccggc ctcctggtgc tggcctccgc catgcggggt tgcccccgag 840
aagccctcca gccgcacctg gcagccatcg ccacagagct ggcacaggcc cacatctgcc 900
aggcatctga aaacgacctc tacctggagc gcctgctgct gtgtgtgcag gctctggtgt 960
ctgtgtgtca tgaggactgt ggcgtggcca gcctgcagct cttggacgtg ctgctgacaa 1020
tagtggccct cgcaggtgct accggcctga gggacaaggc acaggagacg atggactcac 1080
tggccatggt ggagggtgtc agcagctgcc aggacctcta ccgcaagcac attggtcccc 1140
tcctggagcg ggtgaccgcg tcgcaccttg actggaccgc acactcgccg gagctcctgc 1200
agttcagtgt catcgtcgca cagtcaggcc ctgccctggg agaagccctg ccacacgtcg 1260
tgcccacgct gagggcctgt ctgcagccct cccaagaccc gcagatgcgc ctgaagctgt 1320
tctccatcct gtccaccgtg ctgctcagag ccacggacac catcaactcc caggggcagt 1380
ttcccagcta cctcgagacg gtgacaaagg acatcctggc ccccaatctg cagtggcatg 1440
cggggaggac agccgcggcc atccgcacgg ctgccgtgtc ctgcctctgg gcgctcacca 1500
gcagcgaggt cctgtcggca gagcagatac gggacgtgca ggaaacactg atgccccagg 1560
tcctgaccac cctggaggag gattcgaaga tgacgcgact gatctcatgc cgtattatca 1620
acacgttctt aaaaacctcg ggcggcatga cggatccaga gaaactcatc aagatttatc 1680
ctgaactctt aaaacgccta gatgacgtgt ccaacgatgt gaggatggca gccgcctcca 1740

ccttggtcac ctggctgcag tgtgtcaagg gtgccaacgc aaaatcctac tatcagagca 1800
gtgtccagta cctgtaccga gagttgctgg ttcaccttga cgatccagag agggccatcc 1860
aggatgcaat tttagaggtc ctcaaagagg gcagcgggct gttcccagat ctcctggtga 1920
gggagacgga ggccgtcatc cacaagcacc gctcggccac ctactgcgag cagctcctgc 1980
agcatgtgca ggccgtgcca gccacacagt gaccacgctg gtttcagcca cggcacaccc 2040
ttgtccccac ctgagccaga gtttgtggcc tttaaatctc ataaacaagg cacctctgtg 2100
ccagcagtga gactgtgaca gcaagaatgt actcctcagg acacctgccc actctttccc 2160
tggaataaca gcctctgagt ggattctgca tgttatgtga tttgttctgt tcatcgagag 2220
ggctcccaaa catctgcagc tgatttgaaa ttaaaagtaa gtcgcagccg ctcctcccgc 2280
agccacttca gcagcatctt agattttaag cctcacgtgc gcagctggtt catgaactat 2340
tggctgcatc ctgcttaggt gcccaccaag aaggttttta cctacttaac aaaaaagaaa 2400
gaagccaaag tgattagaaa gaaatgaaat ctctttttgg gttctgtcta ctgaaattta 2460
atatctcagt gaacagacta aaaggaattt agaatcctaa caacttacca gatttctcct 2520
gttttaaata tactgggact ttaaaggtta tatgtccggt caccgtatgt tttaagtcgg 2580
tgttaatgct aacagtgttg aaaacaatat ttcatgagat ctaattgtgg ttgcccctat 2640
aggtagcagg aaagtaaagt tgcatttccc tctcgcacat tctacaccca agtgcctaaa 2700
agatctcatt gtaagtgggt agtgttaccg gaagccattg tgttcacacg ggggaaatgc 2760
cgtatatatt tttcaacaaa tattaacgtt tatactttca tgtttgaaaa tttaattaag 2820
aatatttgtt ttaaaaa 2837

<210> 85
<211> 1123
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2079396CB1

<400> 85
taagcttgcg accgccattt ttttttttt ttttttttt tttttactcg tgttgcaatt 60
ttactgtaag tttgaaactg tcaatacaaa aagttacaaa gaattgttta taaagaagaa 120
gaagaaaaaa gcaaacccaa ccaagctcaa gcctagatcc gtatttggat tcagctcgtc 180
tcggggcgga gccaggcgtc acggcctccg gattaaagta tccccccggg gagtgtgctc 240
tgtgaatctg ggtggggagg gcgctcggtg ctgttcccag caacccacca ccctcctcct 300
agtgcttgca agaataggca gggaactcag ctgactgcat cagaacctga gaagctggag 360
gctgaagcca gacaccagcc tctcaggact tggggacact atgagcccac tcagccccac 420
aggtctgaat ctttggggag gggagggttc cagtctgcac tctgcccttg accatcaggg 480
caggggatc acactggcta tcggcatcat ctctagcagt ttctccagcc ccagccctag 540
aatcaggcca agctctcagc actgcgtggg tttgatcttg aggattcttt accaccatcc 600
aggactgggg ggctgccgct catgggtttt gttgttaaga gacagggtct cgctctgtca 660
cccaggctgg agtgcagtgg catgatcacg gctcactgta gccttgacct cccgggctca 720
agcgatcctc cggcctcagc ctcccgagta gctgcgacca caggcctgtg ccagcactcc 780
tggcttgctc gtgctcatct gctatgacgt catgaatccc accagctacg acaacgtcct 840
catcaaggtg aggccggcct cctgggtggt ggccccacat ctgggcccgg ccgcacccct 900
gatcccgcct cctccctcca cagtggttcc ctgaggtcac gcatttctgc cgcgggatcc 960
ccatggtgct catcggctgc aagacagacc tggaatgttc cgccaagttt cgggagaatg 1020
tggaggacgt cttccgggag gccgccaagg tggctctcag cgctctgaag aaggcgcaac 1080
ggcagaagaa gcgccggctc tgcctgctgc tctgacccag ggc 1123

<210> 86
<211> 1549
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 5390115CB1

<400> 86
ttccggggga gcggcgcggc ggcgcgggag tgccacctga ttcactgcac ttggaaccag 60
aagatgttgg gctctttcca gaagtcaaat ccttcctggg aggatgaaga tctgccagct 120
ctgaggattc tggctgggaa gaaaaagagc ctcaagcttt gaaggccatt cccatggtag 180

acagagctga tgtgaaaagt ggccatcctg ttgccccagg cacatgacct tctgaagtga 240
ccaggctgaa agggaagcaa tactcgtgtg atctctcacc ccgtcactca gggtggcgca 300
atcacgactc attggctcac tgcagcctag acctcccagc tggagcaatt ctcctgcctc 360
agccttctga gtagctggga ctacagttgg ttctaaagag tggtgagtca gaagagacgt 420
caggcagcaa gcgacttggg ccatggcctc tgacctagac ttctcacctc cggaggtgcc 480
cgagcccact ttcctggaga acctgctacg gtacggactc ttcctgggag ccatcttcca 540
gctcatctgt gtgctggcca tcatcgtacc cattcccaag tcccacgagg cggaggctga 600
accgtctgag cccagaagtg ctgaggtgac gaggaagccc aaggctgctg ttccttctgt 660
gaacaagagg cccaagaaag agactaagaa gaagcggtag aagaggaggc ctgaggagct 720
gggcgggcag ggagagggtc ttggggacag ccctcctggg aatctacatt gtgttccccc 780
gcattccagg ctcagggtct gaggaggctg tgacgcccta tgaccgcaga gatctagaca 840
gtcgtaacag tccccaggct ccagctgggc aatccaccac ttcctcttcc ttctgcttct 900
gtgacggttt agagtcaagg gggctgaaac acactgtgag catagactgt attaggtttg 960
ttcagaagcc gggtcagctc acagagtcac attttcttgc ttagtcatgt gtccctcctt 1020
gagttgcccc ctccttgtgg gtttacacta catttgggag tcattgtcta atgctgacaa 1080
gcacaccctc tcccattatt tgtgcactac agatctcctg ctgatcagtc acctttgttg 1140
ctgctgtgta gacagagcca ggcctcacct gtttgtttag gccaagatgc catggacatg 1200
cagcgttagt gatcccacta gctgcgacag ccaggcccag aaaatgcctg gcgtgagagc 1260
cagcagacag ccaggccggg gtaggcagtg cctgcttctg ctccatcagg tgcaggggat 1320
ttggctgaag gcgtgcatat ttcctgggca caaacttcct gagcctctga aatgggaggc 1380
tcgtcaattt cagaccaacc tcttttcaac ccatcatagc acgttcaagg tgtgcctttt 1440
acttctacct gtacatcccc catcccttca attctttcat tccctgacca gtgagagggt 1500
tcctgggggа agtatggtga ataaactgac atgcatgctt caaaaaaaa 1549

<210> 87
<211> 4820
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1403326CB1

<400> 87
ctcacactct aacgcgtgta gaatagctgg gaccacagtt tcagctgtga gttcaacatg 60
gaggcaaatc agtgcccccт ggttgtggaa ccatcttacc cagacctggt catcaatgta 120
ggagaagtga ctcttggaga agaaaacaga aaaaagctgc agaaaattca gagagaccaa 180
gagaaggaga gagttatgcg ggctgcatgt gctttattaa actcaggagg aggagtgatt 240
cgaatggcca agaaggttga gcatcccgtg gagatgggac tggatttaga acagtctttg 300
agagagctta ttcagtcttc agatctgcag gctttctttg agaccaagca acaaggaagg 360
tgtttttaca tttttgttaa atcttggagc agtggccctt tccctgaaga tcgctctttc 420
aagccccgcc tttgcagcct cagttcttca ttataccgta gatctgagac ctctgtgcgt 480
tccatggact caagagaggc attctgtttc ctgaagacca aaaggaagcc aaaaatcttg 540
gaagaaggac cttttcacaa aattcacaag ggtgtatacc aagagctccc taactcggat 600
cctgctgacc caaactcgga tcctgctgac ctaattttcc aaaaagacta tcttgaatat 660
ggtgaaatcc tgccttttcc tgagtctcag ttagtagagt ttaaacagtt ctctacaaaa 720
cacttccaag aatatgtaaa aaggacaatt ccagaatacg tccctgcatt tgcaaacact 780
ggaggaggct atcttttтat tggagtggat gataagagta gggaagtcct gggatgtgca 840
aaagaaaatg ttgaccctga ctctttgaga aggaaaatag aacaagccat atacaaacta 900
ccttgtgttc atttttgcca accccaacgc ccgataacct tcacactcaa aattgtggat 960
gtgttaaaaa ggggagagct ctatggctat gcttgcatga tcagagtaaa tcccttctgc 1020
tgtgcagtgt tctcagaagc tcccaattca tggatagtgg aggacaagta cgtctgcagc 1080
ctgacaaccg agaaatgggt aggcatgatg acagacacag atccagatct tctacagttg 1140
tctgaagatt ttgaatgtca gctgagtcta tctagtgggc ctccccttag cagaccagtg 1200
tattctaaga aaggtctgga acacaaagct gatctacaac aacatttatt tccagttcca 1260
ccaggacatt tggaatgtac tccagagtcc ctctggaagg agctgtcttt acagcatgaa 1320
ggactaaagg agttaataca caagcaaatg cgacctttct cccagggaat tgtgatcctc 1380
tctagaagct gggctgtgga cctgaacttg caggagaagc caggagtcat ctgtgatgct 1440
ctgctgatag cacagaacag cacccccatt ctctacacca ttctcaggga gcaggatgca 1500
gagggccagg actactgcac tcgcaccgcc tttactttga agcagaagct agtgaacatg 1560
gggggctaca ccgggaaggt gtgtgtcagg gccaaggtcc tctgcctgag tcctgagagc 1620
agcgcagagg ccttggaggc tgcagtgtct ccgatggatt accctgcgtc ctatagcctt 1680
gcaggcaccc agcacatgga agccctgctg cagtccctcg tgattgtctt actcggcttc 1740

aggtctctct tgagtgacca gctcggctgt gaggttttaa atctgctcac agcccagcag 1800
tatgagatat tctccagaag cctccgcaag aacagagagt tgtttgtcca cggcttacct 1860
ggctcaggga agaccatcat ggccatgaag atcatggaga agatcaggaa tgtgtttcac 1920
tgtgaggcac acagaattct ctacgtttgt gaaaaccagc ctctgaggaa ctttatcagt 1980
gatagaaata tctgccgagc agagacccgg aaaactttcc taagagaaaa ctttgaacac 2040
attcaacaca tcgtcattga cgaagctcag aatttccgta ctgaagatgg ggactggtat 2100
gggaaggcaa aaagcatcac tcggagagca aagggtggcc caggaattct ctggatcttt 2160
ctggattact ttcagaccag ccacttggat tgcagtggcc tccctcctct ctcagaccaa 2220
tatccaagag aagagctcac cagaatagtt cgcaatgcag atccaatagc caagtactta 2280
caaaaagaaa tgcaagtaat tagaagtaat ccttcattta acatccccac tgggtgcctc 2340
gaggtatttc ctgaagccga atggtcccag ggtgttcagg gaaccttacg aattaagaaa 2400
tacttgactg tggagcaaat aatgacctgt gtggcagaca cgtgcaggcg cttctttgat 2460
aggggctatt ctccaaagga tgttgctgtg cttgtcagca ccgcaaaaga agtggagcac 2520
tataagtatg agctcttgaa agcaatgagg aagaaaaggg tggtgcagct cagtgatgca 2580
tgtgatatgt tgggtgatca cattgtgttg gacagtgttc ggcgattctc aggcctggaa 2640
aggagcatag tgtttgggat ccatccaagg acagctgacc cagctatctt acccaatgtt 2700
ctgatctgtc tggcttccag ggcaaaacaa cacctgtata tttttccgtg gggtggccat 2760
taggaagaac tccaaatcaa aatgctatgt aaatgtctat gggtgacagt ctgctgatgg 2820
tagaaacctt tctttttagt tcacaagtca gagatttgga cggagctgac acaaagagtt 2880
tggagctccc ccatttctgg ctctcctttc aggggttcct tccccaaccc ttttcagcag 2940
cggtggctgc cccccattct gacccctgac tcttccagcc agaaagatgg tggttttcta 3000
aaggaacttt agctgtcctg cacaatgccg atctgtgtct tgcattttgg gtaaaagcca 3060
taaaaataag aaactcagcc tgtggccttt ctttcttcca aggctgggct tctttttta 3120
agtgacttca tgcagtttgt tgcttttaaa aatttgtcca gaatcgtttt ctgcagaagc 3180
atggtctgtt aggagcttac tggccgtagc agaagcaatt gtttcctgaa ttcttgacat 3240
ttatctttgc tgtattcatt tagggcttgg gagagtccga agataattca gtcactgtca 3300
gattaataat tttgtcagga caaagaatac cgttatgatt atttaatcct ttaaaattgt 3360
ggtctccaga gcttgttctc agaatggccc agaccaagcc ttaattgtga tagtgaatat 3420
taatggtcac tttaaggaga aattataggc caagatgaaa tgaacataaa cctgtttgcc 3480
ctggctttca gtggaagatg atattagaga ccaaaatctg gttctgaagg tgtgtatcag 3540
ccctaaggtg aaccagactt gggaaagatt gtctttaaaa atcaatgagt ttatgtttta 3600
acttctcagc ttagttctat gcattgctct ataacacacc tagttaagtt ttatgttatt 3660
cttgaactgt gatttttttt ctatttactt tcatggtttg gtgggccatt gttatggact 3720
gaatgtttgt gtcccaccct tcacccccaa attcccgtgt tgaagcccca acctgcactg 3780
tggagctggg gctgctaagg aagtaattaa ggttacatga agtcatggtg gggctctgat 3840
ctgctaaggt tggtgtcctt atagggagag accccagaga gcttgttccc tccctccctg 3900
tgcatgcaaa caagagggca tgggagcaca cagagagatg gcagccacct acaagccaag 3960
aggagaagcc tcacaatcaa actctcgctg ctggcgagag tcttggactc tgtcttggac 4020
ttccagcctc cagactgtga gaaacaaatt tctgttgttt cagcttctca gtctctggtg 4080
ttttgttatt gcagcctgag aacacagctg tacgattatt tgtcaaacag aaaacactga 4140
tacttaacaa tgctaatgca attatttatt tgcttttcag tctctacaaa acgttctaaa 4200
acactaatct aaatattaac agtaaaatat ttgcataact aatggaaact aagaaatcat 4260
atgaccaata tttcacttat tggtaatctt actctactga tttcccccca gactgtgatt 4320
tttggaactt ccttgccttt ctcctgtctt tctgtgttta ttcatggaat tccagttatc 4380
tggggcttga aattgcaggc tctcctaact taagcaaaat ctgacagatc agcaaaatga 4440
gataaatgtt tctttttttct ttctgactgc attaaatcag atacaactca gcattaaaaa 4500
gctatctttg taaatgttgt tactaataaa ttagtcttat aagatccctg gactttggag 4560
ttgttgcaat gtctttgaga gtaattcttt aaaagtctaa tttcgactgg ttgtatctct 4620
ttatgattta ttgccccact aacaacattt gaaacaatat aatattttaa aatgtataaa 4680
taattatgaa tttttgttta gaacaaagag gattactgat atttgtttcc ctatgaatgg 4740
caaaaggttt agcttactac tgcatttctg ttttaaataa aaagttgaga gtttgtgtct 4800
cattaaactg gaaaaaaaaa 4820

<210> 88
<211> 3599
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7690129CB1

<400> 88

ggggccccgt cctccagacc tggctgcagg acctgctgcg tcgtgggctg gtgcgggctg 60
cccagagcac aggagcctgg attgtcactg ggggtctgca cacgggcatc ggccggcatg 120
ttggtgtgcg ctgtacggga ccatcagatg gccagcactg ggggcaccaa ggtggtggcc 180
atgggtgtgg ccccctgggg tgtggtccgg aatagagaca ccctcatcaa ccccaagggc 240
tcgttccctg cgaggtaccg gtggcgcggt gacccggagg acggggtcca gtttcccctg 300
gactacaact actcggcctt cttcctggtg gacgacggca cacacggctg cctgggggggc 360
gagaaccgct tccgcttgcg cctggagtcc tacatctcac agcagaagac gggcgtggga 420
gggactggaa ttgacatccc tgtcctgctc ctcctgattg atggtgatga gaagatgttg 480
acgcgaatag agaacgccac ccaggctcag ctcccatgtc tcctcgtggc tggctcaggg 540
ggagctgcgg actgcctggc ggagaccctg gaagacactc tggccccagg gagtggggga 600
gccaggcaag gcgaagcccg agatcgaatc aggcgtttct ttcccaaagg ggaccttgag 660
gtccttcagg cccaggtgga gaggattatg acccggaagg agctcctgac agtctattct 720
tctgaggatg ggtctgagga attcgagacc atagttttga aggcccttgt gaaggcctgt 780
gggagctcgg aggcctcagc ctacctggat gagctgcgtt tggctgtggc ttggaaccgc 840
gtggacattg cccagagtga actctttcgg ggggacatcc aatggcggtc cttccatctc 900
gaagcttccc tcatggacgc cctgctgaat gaccggcctg agttcgtgcg cttgctcatt 960
tcccacggcc tcagcctggg ccacttcctg accccgatgc gcctggccca actctacagc 1020
gcggcgccct ccaactcgct catccgcaac cttttggacc aggcgtccca cagcgcaggc 1080
accaaagccc cagccctaaa agggggagct gcggagctcc ggccccctga cgtggggcat 1140
gtgctgagga tgctgctggg gaagatgtgc gcgccgaggt acccctccgg gggcgcctgg 1200
gaccctcacc caggccaggg cttcggggag agcatgtatc tgctctcgga caaggccacc 1260
tcgccgctct cgctggatgc tggcctcggg caggcccct ggagcgacct gcttctttgg 1320
gcactgttgc tgaacagggc acagatggcc atgtacttct gggagatggg ttccaatgca 1380
gtttcctcag ctcttggggc ctgtttgctg ctccgggtga tggcacgcct ggagcctgac 1440
gctgaggagg cagcacggag gaaagacctg gcgttcaagt ttgaggggat gggcgttgac 1500
ctctttggcg agtgctatcg cagcagtgag gtgagggctg cccgcctcct cctccgtcgc 1560
tgcccgctct ggggggatgc cacttgcctc cagctggcca tgcaagctga cgcccgtgcc 1620
ttctttgccc aggatggggt acagtctctg ctgacacaga agtggtgggg agatatggcc 1680
agcactacac ccatctgggc cctggttctc gccttctttt gccctccact catctacacc 1740
cgcctcatca ccttcaggaa atcagaagag gagcccacac gggaggagct agagtttgac 1800
atggatagtg tcattaatgg ggaagggcct gtcgggacgg cggacccagc cgagaagacg 1860
ccgctggggg tcccgcgcca gtcgggccgt ccgggttgct gcgggggccg ctgcgggggg 1920
cgccggtgcc tacgccgctg gttccacttc tggggcgcgc cggtgaccat cttcatgggc 1980
aacgtggtca gctacctgct gttcctgctg cttttctcgc gggtgctgct cgtggatttc 2040
cagccggcgc cgcccggctc cctggagctg ctgctctatt tctgggcttt cacgctgctg 2100
tgcgaggaac tgcgccaggg cctgagcgga ggcgggggca gcctcgccag cgggggcccc 2160
gggcctggcc atgcctcact gagccagcgc ctgcgcctct acctcgccga cagctggaac 2220
cagtgcgacc tagtggctct cacctgcttc ctcctgggcg tgggctgccg gctgaccccg 2280
ggtttgtacc acctgggccg cactgtcctc tgcatcgact tcatggtttt cacggtgcgg 2340
ctgcttcaca tcttcacggt caacaaacag ctggggccca agatcgtcat cgtgagcaag 2400
atgatgaagg acgtgttctt cttcctcttc ttcctcggcg tgtggctggt agcctatggc 2460
gtggccacgg aggggctcct gaggccacgg gacagtgact tcccaagtat cctgcgccgc 2520
gtcttctacc gtccctacct gcagatcttc gggcagattc cccaggagga catggacgtg 2580
gccctcatgg agcacagcaa ctgctcgtcg gagcccggct tctgggcaca ccctcctggg 2640
gcccaggcgg gcacctgcgt ctcccagtat gccaactggc tggtggtgct gctcctcgtc 2700
atcttcctgc tcgtggccaa catcctgctg gtcaacttgc tcattgccat gttcagttac 2760
acattcggca aagtacaggg caacagcgat ctctactgga aggcgcagcg ttaccgcctc 2820
atccgggaat tccactctcg gcccgcgctg gccccgccct ttatcgtcat ctcccacttg 2880
cgcctcctgc tcaggcaatt gtgcaggcga ccccggagcc cccagccgtc ctccccggcc 2940
ctcgagcatt tccgggttta cctttctaag gaagccgagc ggaagctgct aacgtgggaa 3000
tcggtgcata aggagaactt tctgctggca cgcgctaggg acaagcggga gagcgactcc 3060
gagcgtctga agcgcacgtc ccagaaggtg gacttggcac tgaaacagct gggacacatc 3120
cgcgagtacg aacagcgcct gaaagtgctg gagcgggagg tccagcagtg tagccgcgtc 3180
ctggggtggg tggccgaggc cctgagccgc tctgccttgc tgcccccagg tgggccgcca 3240
ccccctgacc tgcctgggtc caaagactga gccctgctgg cggacttcaa ggagaagccc 3300
ccacagggga ttttgctcct agagtaaggc tcatctgggc ctcggcccca gcacctggtg 3360
gccttgtcct tgaggtgagc cccatgtcca tctgggccac tgtcaggacc acctttggga 3420
gtgtcatcct tacaaaccac agcatgcccg gctcctccca gaaccagtcc cagcctggga 3480
ggatcaaggc ctggatcccg ggccgttatc catctggagg ctgcagggtc cttggggtaa 3540
cagggaccac agacccctca ccactcacag attcctcaca ctggggaaat aaaccattc 3599

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
27 December 2001 (27.12.2001)

PCT

(10) International Publication Number
WO 01/098353 A3

(51) International Patent Classification[7]: C12N 15/12, C07K 14/47, 16/18, A61K 38/17

(21) International Application Number: PCT/US01/19862

(22) International Filing Date: 20 June 2001 (20.06.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/212,890	20 June 2000 (20.06.2000)	US
60/213,466	23 June 2000 (23.06.2000)	US
60/214,601	27 June 2000 (27.06.2000)	US
60/222,372	31 July 2000 (31.07.2000)	US
60/231,435	8 September 2000 (08.09.2000)	US
60/232,889	15 September 2000 (15.09.2000)	US

(71) Applicant *(for all designated States except US)*: **INCYTE GENOMICS, INC.** [US/US]; 3160 Porter Drive, Palo Alto, CA 94304 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **HILLMAN, Jennifer, L.** [US/US]; 230 Monroe Drive #17, Mountain View, CA 94040 (US). **TANG, Y., Tom** [US/US]; 4230 Ranwick Court, San Jose, CA 95118 (US). **YUE, Henry** [US/US]; 826 Lois Avenue, Sunnyvale, CA 94087 (US). **ELLIOTT, Vicki, S.** [US/US]; 3770 Polton Place Way, San Jose, CA 95121 (US). **TRIBOULEY, Catherine, M.** [FR/US]; 1121 Tennessee Street #5, San Francisco, CA 94107 (US). **LEE, Ernestine, A.** [US/US]; 624 Kains Street, Albany, CA 94706 (US). **RAMKUMAR, Jayalaxmi** [IN/US]; 34359 Maybird Circle, Fremont, CA 94555 (US). **LAL, Preeti** [IN/US]; P.O. Box 5142, Santa Clara, CA 95056 (US). **XU, Yuming** [US/US]; 1739 Walnut Drive, Mountain View, CA 94040 (US). **WARREN, Bridget, A.** [US/US]; 10130 Parwood Drive #2, Cupertino, CA 95014 (US). **HAFALIA, April, J., A.** [US/US]; 2227 Calle de Primavera, Santa Clara, CA 95054 (US). **BAUGHN, Mariah, R.** [US/US]; 14244 Santiago Road, San Leandro, CA 94577 (US). **AZIMZAI, Yalda** [US/US]; 5518 Boulder Canyon Drive, Castro Valley, CA 94552 (US). **BATRA, Sajeev** [US/US]; 555 El Camino Real #709, Sunnyvale, CA 94087 (US). **BURFORD, Neil** [GB/US]; 105 Wildwood Circle, Durham, CT 06422 (US). **YAO, Monique, G.** [US/US]; 111 Frederick Court, Mountain View, CA 94043 (US). **NGUYEN, Danniel, B.** [US/US]; 1403 Ridgewood Drive, San Jose, CA 95118 (US). **LU, Dyung, Aina, M.** [US/US]; 233 Coy Drive, San Jose, CA 95123 (US). **WALIA, Narinder, K.** [US/US]; 890 Davis Street #205, San Leandro, CA 94577 (US). **GANDHI, Ameena, R.** [US/US]; 837 Roble Avenue #1, Menlo Park, CA 94025 (US). **AU-YOUNG, Janice** [US/US]; 233 Golden Eagle Lane, Brisbane, CA 94005 (US). **PATTERSON, Chandra** [US/US]; 490 Sherwood Way #1, Menlo Park, CA 94025 (US).

(74) Agents: **HAMLET-COX, Diana** et al.; Incyte Genomics, Inc., 3160 Porter Drive, Palo Alto, CA 94304 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*

(88) Date of publication of the international search report:
10 October 2002

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 01/098353 A3

(54) Title: SECRETED PROTEINS

(57) Abstract: The invention provides human secreted proteins (SECP) and polynucleotides which identify and encode SECP. The invention also provides expression vectors, host cells, antibodies, agonists, and antagonists. The invention also provides methods for diagnosing, treating, or preventing disorders associated with aberrant expression of SECP.

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 01/19862

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12N15/12 C07K14/47 C07K16/18 A61K38/17

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C12N C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)
BIOSIS, SEQUENCE SEARCH

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
E	WO 01 49728 A (KATO SEISHI ;KIMURA TOMOKO (JP); PROTEGENE INC (JP); SAGAMI CHEM R) 12 July 2001 (2001-07-12) page 558- page 563 page 155 -page 156; claims 1,4 ---	1-4, 6-19,22, 25-45,89
P,X	WO 01 40466 A (STEWART TIMOTHY A ;BAKER KEVIN P (US); DEFORGE LAURA (US); DESNOYE) 7 June 2001 (2001-06-07) * see also AU 2474700 with publication date of 19.06.00 * claims 3,12; figures 195,196 ---	1-4, 6-19,22, 25-45,89
P,X	EP 1 074 617 A (HELIX RES INST) 7 February 2001 (2001-02-07) * SEQ ID NO: 17164 and 17165 * claim 8 -----	1-4, 6-19,22, 25-45,89

☐ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
18 April 2002	16.07.02

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Hillenbrand, G

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 01/19862

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. [X] Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

see FURTHER INFORMATION sheet PCT/ISA/210

2. [X] Claims Nos.: 20-21, 23-24
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

see FURTHER INFORMATION sheet PCT/ISA/210

3. [ ] Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. [ ] As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. [ ] As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. [ ] As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. [X] No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

1-19, 22, 25-44, (all partially), 45, 89

Remark on Protest

[ ] The additional search fees were accompanied by the applicant's protest.

[ ] No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. Claims: 1-19, 22, 25-44 (all partially), 45, 89

Invention 1

The subject-matter of this group of claims is directed to a polypeptide of SEQ ID NO: 1 encoded by the DNA of SEQ ID NO:45, a recombinant polynucleotide comprising said polynucleotide, a cell (transgenic organism) transformed with said polynucleotide, a method for producing said polypeptide, an antibody which binds to said polypeptide, a method for detecting a target polynucleotide in a sample by using said polynucleotide, a composition comprising said polypeptide, use of said composition for treating a disease, screening methods for agonists or antagonists (or compounds which bind) by using said polypeptide, the agonists or antagonists obtained by these methods, screeing methods for compounds that modulate the activity of said polypeptide and screening methods for a compound for effectiveness in altering expression of a target polynucleotide,methods for assessing toxicity of a test compound by using said polynucleotide, and diagnostic tests.

Inventions 2-44

The subject-matter of the residual parts of the claims mentioned above (SEQ ID NOs: 2-44 and 46-88) and claims 46-88 (polypeptides) and claims 90-132 is directed to further inventions 2-44 directed to further polypeptides, their corresponding DNA sequences and their uses as described above.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.1

Although claim 18 is directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.

Although claims 32 and 34 are directed to a diagnostic method practised on the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.

Continuation of Box I.2

Claims Nos.: 20-21, 23-24

Present claims 20-21 and 23-24, which are directed to all possible agonists or antagonists, relate to such a large number of possibly known compounds that a meaningful search was impossible.

The applicant's attention is drawn to the fact that claims, or parts of claims, relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 01/19862

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 0149728 A	12-07-2001	AU 2404001 A	16-07-2001
		WO 0149728 A2	12-07-2001
WO 0140466 A	07-06-2001	AU 2055401 A	12-06-2001
		AU 2192800 A	12-07-2000
		AU 2474700 A	19-06-2000
		AU 2879400 A	31-07-2001
		AU 3514400 A	28-09-2000
		AU 3774300 A	18-12-2000
		AU 5152700 A	02-01-2001
		AU 6391000 A	19-02-2001
		EP 1173563 A1	23-01-2002
		EP 1220905 A2	10-07-2002
		EP 1210418 A1	05-06-2002
		EP 1208195 A2	29-05-2002
		EP 1141289 A2	10-10-2001
		EP 1135495 A2	26-09-2001
		WO 0153486 A1	26-07-2001
		WO 0053758 A2	14-09-2000
		WO 0073454 A1	07-12-2000
		WO 0077037 A2	21-12-2000
		WO 0109327 A2	08-02-2001
		WO 0140466 A2	07-06-2001
		WO 0037640 A2	29-06-2000
		US 2002058309 A1	16-05-2002
		US 2002072496 A1	13-06-2002
		US 2002072067 A1	13-06-2002
		US 2002072092 A1	13-06-2002
		US 2002072497 A1	13-06-2002
		AU 1932000 A	03-07-2000
		AU 3107700 A	28-09-2000
		AU 5460100 A	18-12-2000
		AU 5591100 A	18-12-2000
		EP 1185648 A2	13-03-2002
		EP 1141285 A2	10-10-2001
		EP 1159422 A1	05-12-2001
		WO 0073348 A2	07-12-2000
		WO 0073452 A2	07-12-2000
		WO 0053750 A1	14-09-2000
		AU 2224800 A	28-09-2000
		AU 2883600 A	28-09-2000
		AU 2883700 A	09-01-2001
		AU 5459900 A	18-12-2000
		WO 0053754 A1	14-09-2000
		WO 0053756 A2	14-09-2000
		WO 0078961 A1	28-12-2000
		AU 1749800 A	04-10-2000
		AU 2390700 A	05-02-2001
		AU 2596700 A	28-09-2000
		AU 2883900 A	30-01-2001
		WO 0053753 A2	14-09-2000
		WO 0104311 A1	18-01-2001
		WO 0105836 A1	25-01-2001
EP 1074617 A	07-02-2001	AU 6180800 A	19-02-2001
		AU 6180900 A	19-02-2001
		AU 6181000 A	19-02-2001
		AU 6181100 A	19-02-2001

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 01/19862

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP 1074617 A		AU 6181200 A	19-02-2001
		AU 6181300 A	19-02-2001
		AU 6181400 A	19-02-2001
		AU 6181500 A	19-02-2001
		AU 6181600 A	19-02-2001
		AU 6315800 A	19-02-2001
		EP 1074617 A2	07-02-2001
		EP 1205549 A1	15-05-2002
		WO 0109315 A1	08-02-2001
		WO 0109345 A1	08-02-2001
		WO 0109316 A1	08-02-2001
		WO 0109349 A1	08-02-2001
		WO 0109317 A1	08-02-2001
		WO 0109318 A1	08-02-2001
		WO 0109319 A1	08-02-2001
		WO 0109346 A1	08-02-2001
		WO 0109320 A1	08-02-2001
		WO 0109321 A1	08-02-2001
		WO 0109322 A1	08-02-2001
		WO 0109323 A1	08-02-2001